

Papers in Southeast Asian linguistics No.16

PACIFIC LINGUISTICS

FOUNDING EDITOR: Stephen A. Wurm

EDITORIAL BOARD: Malcolm D. Ross and Darrell T. Tryon (Managing Editors),
Thomas E. Dutton, Nikolaus P. Himmelmann, Andrew K.
Pawley

Pacific Linguistics is a publisher specialising in linguistic descriptions, dictionaries, atlases and other material on languages of the Pacific, the Philippines, Indonesia and southeast Asia. The authors and editors of Pacific Linguistics publications are drawn from a wide range of institutions around the world.

Pacific Linguistics is associated with the Research School of Pacific and Asian Studies at the Australian National University. Pacific Linguistics was established in 1963 through an initial grant from the Hunter Douglas Fund. It is a non-profit-making body financed largely from the sales of its books to libraries and individuals throughout the world, with some assistance from the School.

The Editorial Board of Pacific Linguistics is made up of the academic staff of the School's Department of Linguistics. The Board also appoints a body of editorial advisors drawn from the international community of linguists. Publications in Series A, B and C and textbooks in Series D are refereed by scholars with relevant expertise who are normally not members of the editorial board.

To date Pacific Linguistics has published over 400 volumes in four series:

- **Series A: Occasional Papers;** collections of shorter papers, usually on a single topic or area.
- **Series B: Monographs** of intermediate length.
- **Series C: Books;** publications of greater length, especially reference books such as dictionaries and grammars, and conference proceedings.
- **Series D: Special Publications;** including archival materials, pedagogical works, maps, audiovisual productions, and materials that do not fit into the other series.

PAPERS IN SOUTHEAST ASIAN LINGUISTICS No.16

edited by
M. Clark



Pacific Linguistics

Research School of Pacific and Asian Studies
The Australian National University
Canberra

Published by Pacific Linguistics
Research School of Pacific and Asian Studies
The Australian National University
PO Box 1428
Canberra ACT 2601
Australia

Copyright © The several authors

First published 1997

ISSN 0078-7531, 0078-9178
ISBN 0 85883 465 9

Copiedited by Basil Wilson, typeset by Sue Tys, map work by Kay Dancey

Printed by ANU Printing Service, Canberra
Bound by F & M Perfect Bookbinding, Canberra

TABLE OF CONTENTS

EDMUND A. ANDERSON	
The use of speech levels in Sundanese	1–45
FERDINAND DE HAAN	
Khmer and the theory of modality	47–66
JUDY HO	
Socio-semantic aspects of human measure words in Cantonese	67–77
KITIMA INDRAMBARYA	
The status of the word <i>hây</i> in Thai	79–118
CHARLES PAUS	
Variability in Cambodian copular constructions: a semantic analysis	119–131

THE USE OF SPEECH LEVELS IN SUNDANESE

EDMUND A. ANDERSON

1. INTRODUCTION

The use of lexical variables to represent social meaning has been commented on for various languages, most notably Javanese, Japanese and Korean. An update which does justice to these studies, even of the Javanese material only, is beyond the scope of this article, but the reader is referred to the monograph by Errington (1985), which appears to be the most recent treatment of Javanese.

Sundanese speech levels have been treated by Kern (1906), Kats and Soeriadiradja (1927), Eringa (1949), Satjadibrata (1956), Wirakusumah and Djajawiguna (1957), Noorduyn (1963), Wessing (1974), Djajawiguna (1978), Ayatrohaedi (1980), and Soedradjat (1986).

Although Sundanese speech levels have been discussed by a number of scholars, these studies have not been based on the analysis of actual usage, but of reported usage. This has yielded a standard model which, in the words of Wessing (1974:12), appeals to “features of the social environment in which the speech event is taking place”, most notably, “a) social status of the addressee or referent, b) social status of the speaker, c) the difference (if any) between a) and b) [and], d) the degree of friendship (intimacy) between the speaker and the addressee”.

This paper presents this model in some detail, then discusses the results of the analysis of 60 or so texts to observe how speech levels are actually used by Sundanese interlocutors. Results of this analysis reveal previously unreported aspects of interlocutors’ knowledge governing use of speech levels.

2. THE SPEECH SITUATION

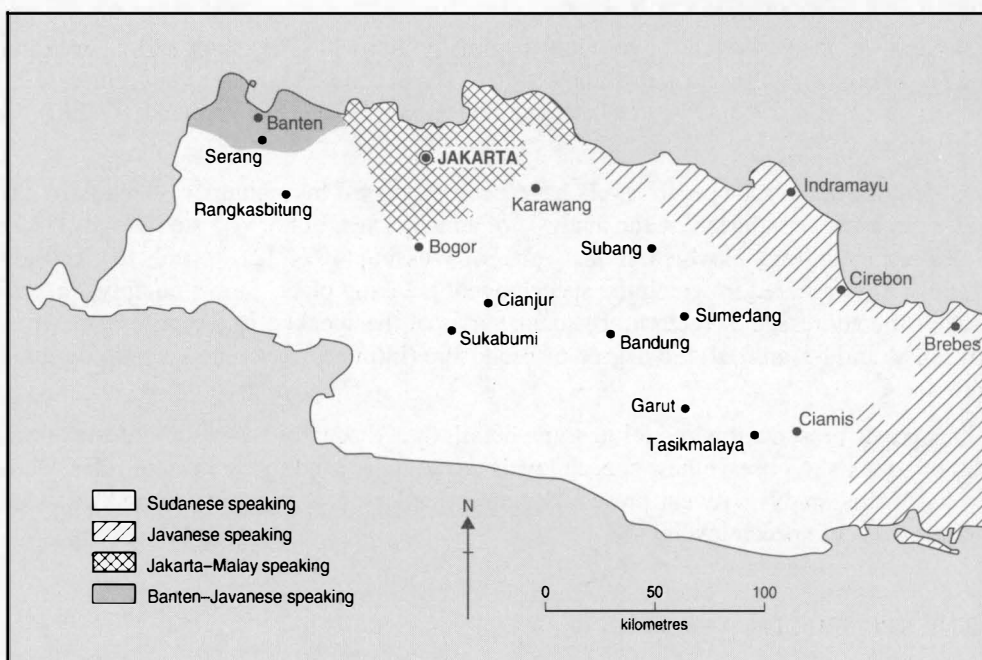
The estimated 1994 population of Indonesia is about 200 million, of which over 118 million reside on the island of Java. Of those 118 million, an estimated 28 million are speakers of Sundanese, the regional language of West Java.¹

The Sundanese language situation is succinctly summed on the macro-level by Harsojo (1983:300–301):

Nowadays Sundanese is used widely among the population of West Java. In villages, the language of instruction is Sundanese, whereas, in towns,

¹ Asian Net – Indonesian Home Page, www.asiannet.com/indonesia/indonesia.

Sundanese is utilized primarily in the family circle, in conversation among friends and intimate acquaintances, and also in public and official places between people who are aware they both know Sundanese. **With regard to language refinement, it is often said, that pure and refined Sundanese is to be found in the area of Priangan, that is, in the regencies of Ciamis, Tasikmalaya, Garut, Bandung, Sumedang, Sukabumi and Cianjur.** Even now, the Cianjur dialect is still considered the most refined Sundanese. From Cianjur came the songs for lute and flute referred to as *Cianjuran*. Considered less refined is the Sundanese near the north coast of Java, for example, that spoken in Banten, Karawang, Bogor and Cirebon. [Furthermore,] The language of the Baduy, which is spoken in south Banten, is archaic Sundanese.



MAP OF RELEVANT PARTS OF WEST AND CENTRAL JAVA

(After R.R. Hardjadibrata (1985), *Sundanese: a syntactical analysis*, p.2. PL, D-65.)

The Baduy people have fascinated other Indonesians and foreigners alike for some time now. A subgroup of Sundanese people, the Baduy have lived apart from Islamic and Western influences with which most Sundanese people have been interacting for centuries. Unlike other Sundanese they have resisted cultural change. They did not convert to Islam. The Baduy reside in the mountains at the extreme western end of Java, in south Banten. Here they remained out of reach of imperial and Islamic cultural incursions. One story has it that their ancestors, defeated warriors of the pre-Islamic Kingdom of Pajajaran, fled there after suffering defeat in battle at the hands of the Islamic kingdom of Banten, probably in

the late fifteenth or early sixteenth century, but this is now believed to be a romanticised view.²

In modern times, the psychological impact of meeting some Baduy people has been recorded by the Sundanese literary scholar, Ajip Rosidi, who highlighted their social characteristics which may be close to those of pre-contact Sundanese people. Included in these characteristics are, of course, those of their language. Rosidi (1980:128) writes:

I have never done a formal study, but when meeting several of the Baduy, I noted no unique physical characteristics. There were, however, striking contrasts to other Sundanese people. Immediately evident were features developed as a response to their environment, to nature and as a result of their education – in short, culture traits: the distinctive black or dark blue clothing, the ancient head scarf, the behavior so full of self-confidence [as opposed to the shyness, *malu*, of other Sundanese], and 4) **the language, which does not symbolize relative social levels of speaker and addressee.**

Speculation about how speech levels came to be an integral part of Sundanese in the areas of West Java has yielded various explanations. The following appears to be the most credible, given geographical factors:

Aside from an emotional, literary evaluation, the existence of the distinction between refined and less refined, and pure and less pure Sundanese may perhaps be explained from the point of view of Sundanese history: **Priangan, for example, was known to have been culturally influenced by the Islamic [Javanese] Kingdom of Mataram. In the 19th century, there were familial and cultural relations between Sundanese nobility, specifically in the area of Sumedang, with [Javanese] nobility in Solo and Yogyakarta. In addition, it is possible that the psychological climate and environment exerted an influence upon certain aspects of language.** (Harsojo 1983:301)

The location of Sumedang in the easternmost part of West Java, along with Ciamis, Tasikmalaya, and Garut (that is, close to what is today Central Java) gives credence to this theory.

On the personal level of interlocutors in actual conversations, historically, Satjadibrata reports (quoted in Soedradjat 1986:108):

The higher speech level was employed among members of the regents' [= chief district administrators'] families. In fact, only upper class people, descendants of the royal families, knew the speech level system. The speech levels were formally taught in [Dutch] schools only in the early 20th century. Speech levels thus became a part of the life of educated Sundanese people (Rosidi 1980).

Indonesian independence in 1949 exerted new influences on language development. According to Soedradjat (1986:108):

The use of speech levels has changed over time in such a way that some levels are now used differently from the way they were used very much earlier. Some of these levels [i.e. P (medium) and LP (very high level)] are rarely used nowadays. Only in the *wayang* [puppet drama] performance is the high level now used [in the way that it was used socially in the past]. In the past, the high level was used to show respect to the *menak* 'aristocrats'. Nowadays, however,

² Prof. Noorduyt has called my attention to the challenging of this interpretation by Bakels (1989).

the distinction in class based on blood is not significant, and therefore the high level [i.e. L] is used to any addressee whom a speaker thinks should be given respect.

Variant terminological systems and models of speech level use exist. Of these, Satjadibrata's (1956) seems to be somewhat of a standard formulation, because Noorduyin, Wessing and Soedradjat rely heavily on it.

According to Satjadibrata (1956:11), two levels constitute the frame for the system: *Kasar* (K) 'low level' or 'general conversational speech' and *Lemes* (L) 'high level' or 'speech in which polite (i.e. *Lemes*) vocabulary is used'.

People are said to be 'speaking *Lemes*' or 'speaking *Kasar*' according to the overall impression the listener derives from the utterances s/he hears.

Finer distinctions are possible within this framework, though very infrequently used: *Lemes Pisan* (LP) 'very polite', *Panengah* (P) 'rather polite', and *Kasar Pisan* (KP) 'crude, earthy'.

3. SOCIAL DETERMINANTS OF SPEECH LEVEL USE

The nature of the *social relationship of interlocutors* influences selection of speech level. Social relationship is a function of the interaction of the relative statuses of interlocutors in the conversation.

Types of social relationships can be seen against the backdrop of social groupings in Sundanese society. Harsojo (1983:305) writes:

Economics, politics and modern ideology, governmental administration, communications, and education have created an **upper social stratum**, consisting of village administrators, teachers, information specialists, office workers, students, members of the armed forces, merchants and entrepreneurs, all of whom possess an outward looking orientation. On the other hand, there is a **lower stratum**, farmers, whose number is great, most of whom are still illiterate and whose life style is still traditional. People on the upper stratum possess economic skills based on the principle of seeking profit and possess connections with middlemen and large merchants in cities. It may also be said that all economic power of the village is centered in the upper stratum, and, generally, the bond between the upper and lower classes takes the form of debt or contracts which do not benefit the lower stratum, whose economy is weak. However, whenever we investigate in West Java, of course, not all villages have experienced the same changes.

Surjadi (1974:22-24) also provides an insight into the social structure of Sundanese people, in the context of modernisation:

Hildred Geertz (Geertz 1963:16-18) made a connection between means of livelihood with the social system. In cities she proposed a social composition consisting of "*the urban elite, the urban middle class and the urban proletariat*." The **urban elite** consists of the diplomatic community and businessmen, communities of foreign businessmen from China, Arab countries, and India. Next, "**an Indonesian metropolitan superculture**" is in the process of forming itself by cultivating symbols such as higher education, ability to speak foreign languages, overseas experience and possession of western produced luxury goods such as automobiles.

The **urban middle class** consists of mid-level civil service employees and other officials such as health officers, teachers and mid-level members of the armed forces. It also includes the groups of skilled laborers: tailors, stone masons, blacksmiths, merchants, electricians, drivers, motorized vehicle mechanics, etc.

Finally, there is the **urban proletariat** consisting of laborers, messengers or household servants, peddlers, pedicab drivers, etc. They generally are unskilled and illiterate.

As for the **village**, Geertz did not propose a classification or even a statement of social composition. She only wrote "most villages are fairly homogeneous both in economic condition and in general outlook."

With regard to the social composition or even social strata in cities, as proposed above, the major part is valid for cities in West Java, especially Bandung. Indeed, in cities in West Java there is no diplomatic community, nor are there foreign businessmen. Even in Bandung the diplomatic community is very small. It includes British citizens at the British Council. The community of foreign businessmen, indeed exists, as, for example, the Japanese.

Nowadays, the **service sector**, especially the field of **entertainment**, is developing in large cities such as Bandung, with the birth of night clubs and bars. Another fast growing area is **transportation**. The number of intercity vehicles in West Java, and also those within the city of Bandung itself, where new routes for motorcycle taxis, honda taxis, etc., are very rapidly being developed compared with previous years...The result of this development for the labor sector is an increase of drivers, conductors, and also passenger agents.

The *social relationship* of the interlocutors affects their choice of terms to **address** and **refer to** each other. The frame of reference within which this adjustment of their speech takes place is that of 'familiarity', or *kekeluargaan* in Indonesian. Surjadi (1974:134) illustrates this principle.

In everyday social interaction, Sundanese people, when meeting a person for the first time, if after they reveal their genealogy (*panca kaki*) and determine there is no family relationship whatsoever, **they then determine each other's positions according to age**. So the younger calls the elder *akang* 'elder brother' [*euceu* 'elder sister'], and in return the elder addresses the younger as *ayi/adi* 'younger brother[/sister]'. The term of address *mang*, which originated from the word 'uncle', besides its use for family relationship, is also frequently utilized for the meaning which has no bearing on family relationship, for example, a person with whom he is not acquainted, but whose status is considered lower, i.e. toward pedicab drivers, ox cart drivers, and so forth.

Furthermore, the interlocutors' *social relationship* also affects their choice of lexical terms in any way referring to themselves or each other, or any other persons whom they may happen to refer to (e.g. **my, your** or **his/her house**).

Figure 1 is a summary of norms for the selection of speech level with reference to the *social relationship* of the interlocutors:

1. STRANGERS speak *Lemes* with each other.
2. a. Lower Status persons speak *Lemes* up to Higher Status persons and
b. Higher Status persons speak *Kasar* down to Lower Status persons.
3. a. A speaker uses *Lemes* to refer to a referent of Higher Status than either speaker/ listener.
b. A speaker uses *Kasar* to refer to a referent of Lower Status than either speaker/ listener.
4. Equal Status persons speak *Lemes* with each other, with 2 exceptions:
a. *Well-acquainted* Equal Status persons speak *Kasar* with each other. and
b. 1) *Younger* Equal Status persons speak *Lemes* to *older* persons while
2) *Older* Equal Status persons speak *Kasar* to *younger* persons.
5. a. A speaker uses *Kasar* to refer to a referent of Same Status who is a *close friend*.
b. A speaker uses *Lemes* to refer to *referent* of Same Status when already using *Lemes* [1, 2a, 4b1)].
c. A speaker uses *Kasar* to refer to a *referent* of Same Status when already using *Kasar* [4a, 4b2)]. while
d. An older speaker uses *Kasar* to refer to a younger referent of Same Status.
6. *Intimate friends* speak *Kasar* with each other.
7. *Some people* (i.e. *lower class*) speak *Kasar* among themselves.
8. A *speaker* uses *Kasar* to refer to himself/herself when *speaking* to intimate friends.

FIGURE 1: SPEECH-LEVEL NORMS: ADDRESS/SECOND-PERSON REFERENCE AND FIRST-PERSON REFERENCE (After Satjadibrata 1956)

This arrangement of choice features shows that *social relationship* is the framework within which the system operates. STRANGERS, at one end of the scale, are presumed *always* to speak *Lemes* with each other, while *intimate friends*, at the other end, always speak *Kasar*.

In this model, the people whom Satjadibrata mentions in category 7 do not seem to be active interlocutors in the speech-level system, and we are not told anything more about their ability in this area.

All others are governed by interlocutors' social relationship, that is their *relative statuses*. Only when interlocutors' *statuses* are equal will *intimacy* and *relative age* exert an influence on choice of speech level.

4. SPEECH LEVEL VARIANT TYPES

A speaker must know four patterns of lexical variables to convey the appropriate attitude, among which are *reserve* (with strangers), *respect* and/or *humility* (to one's betters) or *solidarity* (with one's intimates). The four patterns are as follows:

ATTITUDE	Type I Address and Reference	Type II Address and Reference	Type III Address and Reference	Type IV Non-Address and Non-Reference 'neutral polite'*
<i>Respect</i> 'Respectful words'*	Lr <i>ningali</i> look at		Lr <i>rambut</i> hair	LE <i>dugi</i> until
POLITE	<i>bumi</i> house			<i>énjing</i> tomorrow
	<i>ibu</i> mother	Lr/Lh <i>éwéd</i> confused		as in... <i>Dugi ka énjing!</i> Till tomorrow!
<i>Humility/ Respect</i> 'Humble words'*	Lh <i>ningal</i> look at			
	<i>rorompok</i> house			
POLITE	<i>biang</i> mother		Lh/K <i>buuk</i> hair	
UNMARKED	K <i>nénjo</i> look at	K <i>bingung</i> confused		K <i>datang</i> until
	<i>imah</i> house			<i>isuk</i> tomorrow
	<i>indung</i> mother			as in ... <i>Datang ka isuk.</i> Till tomorrow.

FIGURE 2: PATTERNING OF VARIANTS
(After Satjadibrata 1956) [*Follows Noorduy'n's (1963) English terminology]

Lr = *Lemes of respect* (Terms referring to objects, places, actions, etc. associated with the addressee or with someone being referred to whom the speaker wishes to respect.)

Lh = *Lemes of humility* (Terms referring to objects, places, actions, etc. associated with a speaker, who wishes by humbling himself to show respect to the addressee or person being referred to.)

LE = *General Lemes* (Terms not associated with speaker or addressee, but nevertheless are considered to be *Lemes*.)

K = *Kasar* terms.

In Figure 2, the **Type I** variable has a different word or form of the same word for *Lemes of Respect*, *Lemes of Humility* and *Kasar*. A speaker has two means of showing respect toward the addressee or a third person:

1. By referring to the addressee's or a third person's activity (e.g. 'looking at'), with a *respectful word* (Lr).

Bapa ningali ka abdi. (Lr) You, sir are looking at me.

2. Or, by referring to his/her own activity with a *humble word* (Lh). By humbling himself/herself, a speaker indicates respect for addressees or third persons.

Abdi ningal ka Bapa. (Lh) I am looking at you.

Both variants are *Lemes* and are forms of the same word. The *Kasar* variant is a different word, *néngo*.

Different verb forms operate similarly:

Bapa ditingal ku abdi. (Lh) You were seen by me.

The speaker refers to his/her own action with a *humble word* (Lh). In so doing, s/he symbolises respect for the addressee.

Buku éta ditingali ku Bapa. (Lr) That book was seen by you.

For the addressee's activity, the speaker chooses a *respectful word* (Lr).

Additionally, there are a few Type I nouns, as well:

rorompok abdi. (Lh) house my (my house)

but,

bumi Bapa. (Lr) house your (your house)

Both variants are different words, here, though both are *Lemes*. The *Kasar* variant is yet another word, *imah*.

Another Type I noun is 'mother': Lh = (*pun*) *biang*, Lr = *ibu* and K = *indung*. All three variants are different words.

Type II variables in Figure 2 have the same word for *Lemes of respect* and *Lemes of humility* and a different word for the *Kasar*. A speaker may respect the addressee or personal referent by referring to a characteristic of either the addressee, referent or himself/herself with the same *Lemes* term.

Abdi éwéd. (Lh) I (am) confused.

Gamparan éwéd. (Lr) You (are) confused.

The *Kasar* term is a different word entirely, *bingung*.

It is unclear why there is no distinction between *respectful words* and *humble words*.³

Type III variables in Figure 2 have one word for *Lemes of respect* and another for *Lemes humble* and *Kasar*. A speaker humbles himself/herself by referring to himself/herself with the same word as is used in *Kasar* speech. That is, the same word would be used in a situation requiring a *Kasar* term symbolising different social status/rank and also solidarity.

³ Prof Noorduynd feels that the collapse of this distinction is diachronically motivated.

rambut Bapa (Lr) *hair* your

buuk abdi (Lh/(K)) *hair* my

A subclass of Type III terms, including *adi* 'younger sibling', *anak* 'child', and *bapa* 'father', make the distinction between *Kasar* and *Lemes* by prepositioning the particle *pun*. Prepositioned *pun* makes the expression *Lemes*: *pun anak* 'my child', *pun bapa* 'my father', both of which are *Lemes* *humble* words (Lh).

Type IV variables, because they are not used for address, and first, second or third person reference have only a *Lemes* and a *Kasar* term. Interlocutors can show respect to each other by choosing *neutral polite* (LE) terms. *Neutral polite* terms can be adjectives, adverbs, nouns, verbs, conjunctions and augmentatives (e.g. *pisan* 'very').

Datang ka isuk. [*Kasar*]

Dugi ka énjing. [*Lemes*]

Until tomorrow! (lit. arrive at tomorrow!)

Mention of speakers' ways of referring to themselves are rather scattered in previous treatments. Item 8 in Figure 1 covers one particular case. Satjadibrata does not mention self-reference for STRANGERS. Presumably they should use *Lemes* to refer to themselves, but I have observed that STRANGERS make very few references to themselves, each other, or, for that matter, to third persons. Furthermore, it appears that Satjadibrata leaves many other contingencies for self-reference unaccounted for. How, for example, does a speaker refer to himself/herself when the addressee is neither an intimate friend, nor someone to whom one speaks in the *Lemes* style? We are not told.

With reference to Figure 1, a further use of *Kasar* relates to scholarly written Sundanese:

In addition, when writing for the general public, *Kasar* is usually used. For scholarly books on mathematics, linguistics, or other scholarly fields, for ease of comprehension, it is better to use *Kasar*.
(Satjadibrata 1956:14)

With regard to the choice of *Lemes* style, choice of vocabulary is not the only aspect to be known by the speaker:

In the *Kasar* level, the manner of speaking plays an important role. Even if the vocabulary is perfect, if the speed and loudness are not proper, the speaker may be regarded as not respecting the addressee.
(Soedradjat 1986:112)

Socially, use of speech levels is one aspect of politeness, which ranges from *Lemes*, 'refined, in accordance with customary law' to *Kasar*, 'lacking refinement'. The *Kasar-Lemes* scale is a theme underlying many aspects of Sundanese culture, from deportment, dress, body movement and so on.

In conclusion, by following norms for use of speech levels, desired social meanings are transmitted. The use of *Kasar* style not only can express awareness of the addressee's lower status (or lack of respect if the norm is ignored) but, in the right circumstances, can symbolise solidarity.

5. SPEECH-LEVEL USE IN DISCOURSE: A TEXT-BASED STUDY

More than sixty texts were elicited from a total of eight different people over a period of two months. Five people were language instructors in Indonesian and Sundanese at a language school in Bandung. One was a university graduate with the Doctorandus degree

in economics, who is now employed in the Department of Taxation in Jakarta. The other four were high-school graduates. The remaining three speakers were village people: two women working as household helpers, a man working as a warehouseman.

The logic of using elicited texts (rather than constructed texts, or interviews of various types) is that the elicitation process focuses interlocutors' attention on the task at hand, whereas the other techniques tap interlocutors' explicit knowledge formulations (of speech-level phenomena, in this case). Elicitation taps interlocutors' tacit knowledge. That is, they will be able to judge 'what sounds appropriate' in a particular situation, but will not be able to explain why it is appropriate.

In these sessions, I described situations in which a speech event might take place. Participants assumed various roles in these speech events. Speech-event enactments were tape-recorded. Each enactment was immediately replayed to the actors and other Sundanese-speaking onlookers, as a check on its authenticity. If a performance as a whole, or any interlocutor's performance, was deemed inauthentic, the performance was repeated (several times, on occasion) until interlocutors and onlookers were satisfied. Recordings were archived and are the basis for this study.

Speech-level terms in texts were identified according to Satjadibrata's (1956) listing and tagged by type: *Lemes Pisan*, *Lemes*, *Panengah*, *Kasar* (but only if choice of other variants was possible, seeing that *Kasar* is the unmarked variant) and *Kasar Pisan*.

For a broad overview, frequencies of occurrence of speech-level terms were tabulated. Numeric values were assigned to each speech level to quantify an impressionistic politeness scale: *Lemes Pisan* (+2), *Lemes* (+1), *Panengah* (+0.5), *Kasar* (0) and *Kasar Pisan* (-1). An average of values for each dyad (i.e. a pair of interlocutors) [Ave.(Dyad)] and for each individual was calculated for all texts using the following formula:

$$M = \frac{\Sigma(2a+b+.75c-e)}{N} \left[\begin{array}{l} \text{Where a, b, c, d and e are the total of } \textit{Lemes Pisan}, \\ \textit{Lemes}, \textit{Panengah}, \textit{Kasar} \text{ and } \textit{Kasar Pisan} \text{ terms,} \\ \text{respectively, in each text, and where N equals the} \\ \text{sum of a, b, c, d and e.} \end{array} \right]$$

This array of Ave.(Dyad)s represents a global view of speech-level use by dyad and by interlocutor across texts.

Dyads with ten or fewer terms of speech-level terms were dropped, because averages fluctuate widely with small numbers of items.

6. RESULTS

The framework for presenting data was a range of *degrees of acquaintance* (i.e. from STRANGERS, to ACQUAINTANCES, CO-WORKERS, FRIENDS, and FAMILY MEMBERS). Averages for each dyad were sorted by degree of *acquaintance*, and arrayed in descending order.

Ranges of Ave.(Dyad)s for the various *degrees of acquaintance* of interlocutors are as follows:

TABLE 1: RANGES OF AVE.(DYAD) AND OVERALL AVE. BY DEGREES OF RELATEDNESS

Relatedness	Overall Ave.	Range of Ave.(Dyad)
Strangers	0.97	0.80–1.17
Acquaintances	0.85	0.21–1.06
Co-Workers	0.85	0.40–1.00
Friends	0.77	0.18–1.00
Family	0.75	0.09–1.00

A closer look reveals that the upper bounds are similar while the lower bounds have great variability. The lower bound of each range appears to distinguish them from one another. One initially confusing fact is the similarity of upper bounds. What is the reason for this similarity? The answer appears to be that, in every degree of acquaintance group, there is a wide variability in the use of speech levels and it will be necessary to investigate further to discover what the reasons for this variability are.

i) STRANGERS⁴ (Refer to Appendix 1)

STRANGERS predictably use *Lemes* style overwhelmingly [Overall Ave.(STRANGERS) = 0.97; $n^5 = 591$].

The following text excerpts were produced by STRANGERS. The venue indicated was the train station in Bandung. A 28-year-old businessman (D) has returned from Jakarta on the evening train and been met by his brother. Leaving the station, they encounter a woman in her forties (S) who seems apprehensive. She approaches and asks directions to Hotel Homan. *Lemes* words are in bold roman type.

- 1 D: *Oh, Ibu badé ka mana? Katingalna* linglung pisan.*
oh madam wants.to to where apparently confused very
Oh, where do you want to go? You seem very confused.
**Tingal* 'see, appear', *ka...na* ADV.
- 2 S: *Leres, Dén.*
Correct, prince/honoured sir.
- 3 *Ibu téh badé milarian panginepan, nyaéta Hotél Homan.*
ma'am FOC wants.to find lodging that.is Hotel Homan
As for me, I'm looking for my lodging, Hotel Homan, that is.
- 4 *Namung Ibu teu acan terang, kumargi Ibu jalmi énggal*
but madam not.yet know because madam person new
- 5 *kitu ti Tasik dongkap ka dieu, namung milarian*
like.that from Tasikmalaya come to here but look.for
- 6 *panginepan teu acan pendak. Margi teu terang.*
lodging not yet reached because not know

4 Abbreviations used in these texts are: ADV = adverbialiser, AUG = augmentative, CAUS = causative, DEF = definite article, FOC = focus, NOM = nominaliser, PASS = passive, REL = relative.

5 That is, the total number of speech-level terms for that degree of relatedness.

But, I'm not sure yet, because I'm new here, the story is, I came here from Tasikmalaya, but I haven't succeeded in finding lodging. Because I don't know Bandung.

- 7 D: *Ibu téh nembé ka Bandung, Ibu téh?*
 madam FOC only.just to Bandung madam FOC
 You, Madam, have just come to Bandung for the first time yourself?

- 11 S: **Leres.**
 Correct.

And so the conversation continues.

The general pattern was to use *Lemes* style almost exclusively among STRANGERS. There were, however, 23 instances where STRANGERS utilised *Kasar* terms. Of these, 10 were used only after interlocutors had been introduced. After being introduced, they were no longer STRANGERS, but new ACQUAINTANCES. These, then, strictly speaking, are *not* instances of STRANGERS using *Kasar* speech level. If these 10 *Kasar* terms are excluded, the Overall Ave.(STRANGERS) rises to 0.99.

Furthermore, of these 23 seemingly misplaced *Kasar* terms in STRANGER dyads, it turns out that 19 are Type IV *KASAR* variants, that is *Lemes Enteng* ('neutral polite') terms.

Since *Lemes Enteng* terms are 'general, non-specific' terms not oriented toward addressee, speaker, or any third person, the choice of Type IV *Kasar* terms must signal something else. In this case, I believe it signals a slight general relaxation of formality.

Three of the remaining four *Kasar* terms are used by a speaker who wishes to show respect by humbling herself, after having conversed with a potential employer.

Finally, social characteristics of interlocutors (i.e. sex/relative age/status difference) as well as various locations of speech events are dispersed throughout the array of STRANGER dyads. This indicates that these are less significant than *degree of acquaintance*. In conclusion, then, texts for STRANGERS are by and large explained by Satjadibrata's model.

ii) ACQUAINTANCES AND CO-WORKERS (Refer to Appendices 2-4)

ACQUAINTANCES and CO-WORKERS overwhelmingly use *Lemes* style to interact with each other [Overall Ave.(ACQUAINTANCES) = 0.85, $n^5 = 341$; Overall Ave.(CO-WORKERS) = 0.85, $n = 651$].

Though the range of averages for ACQUAINTANCE dyads overlaps with that for STRANGERS, the difference between the overall averages for STRANGER and for ACQUAINTANCE dyads is statistically significant:

Overall Ave.(STRANGERS) = 0.97 < 0.85 = Overall Ave.(ACQUAINTANCES)
 ($\chi^2 = 15.33$, $df=2$; $p<.001$)

And it follows that the overall average for CO-WORKER dyads is also statistically different from that for STRANGER dyads.

Since the overall average for ACQUAINTANCE dyads and that for CO-WORKER dyads are the same, these two types of dyads have been unified into a single table and treated as a single category, though identities of individual dyads have been retained.

The range of averages for ACQUAINTANCE and CO-WORKER dyads is wider than that for STRANGERS: for STRANGERS it is 0.80–1.17; for ACQUAINTANCES and CO-WORKERS it is 0.21–1.06.

Twenty-five of 35 (71%) of averages for ACQUAINTANCE and CO-WORKER dyads exceed 0.80 (the lowest average for any STRANGER dyad). That is, 71% of ACQUAINTANCE and CO-WORKER dyads are similar to STRANGER dyads in frequency with which interlocutors use *Lemes* terms.

ACQUAINTANCES and CO-WORKERS use *Lemes* style frequently when interacting. But, they use *Kasar* terms more frequently than do STRANGER. As for STRANGERS, the majority of *Kasar* terms used (64%; 96/150) are chosen in preference to *Lemes Enteng* (Type IV ‘neutral polite’) words, indicating an *overall gradual relaxation of formality* in those social situations.

The remaining 36% of *Kasar* terms consist of those chosen instead of *Lemes respect* or *Lemes humble* terms. This means that interlocutors intend to signal their perceptions of their own, addressees’, and referents’ relative social statuses (Same Status). If we eliminate *Kasar* terms used instead of *Lemes Enteng* variants (i.e. *Kasar*<*Lemes enteng*) from consideration, we are left with *Kasar* terms replacing *Lemes respect* and *Lemes humble* terms (i.e. *Kasar*<*Lemes respect/humble*), and from these we have a clear indication of the social meanings being exchanged (see Appendix 3).

Patterns of *Kasar*<*Lemes respect/humble* for interlocutors in each dyad indicate shared social attitudes toward established social roles (see Appendix 4).

Two major patterns emerge. In Pattern 1, one of the two interlocutors has no *Kasar*<*Lemes respect/humble* terms while the other speaker has between 1 and 15. The average for the second interlocutor is 2.7. In Pattern 2, both interlocutors have 1 or more *Kasar*<*Lemes respect/humble* terms.

Pattern 1: The differences between *Kasar*<*Lemes respect/humble* for the interlocutors are, respectively, 15 (one dyad, 48.2); 4 (one dyad, 49a.1); 3 (one dyad, 44c.2); 2 (two dyads, 44b.2, 49c.2); and 1 (seven dyads, 51a.1, 57a.1, 31a.1, 32.2, 44a.2, 60.2, 49a.2). (See note to Appendix 1.)

Pattern 2: The differences are, respectively, 3 (one dyad, 44a.1); 1 (two dyads, 45.2 and 50.2); and 0 (one dyad, 45a.1).

To interpret these patterns, texts were examined for the social values *showing respect* and/or *humbling oneself*.

In text 48, two women friends, A and B, converse in *Kasar* style, in accordance with the received model (Ref. Figure 1–4a).

In part 2 of the same conversation, C, the nephew of A, enters. Averages for B and C as they interact are:

$$\text{Ave.}(B) = 7 < 100 = \text{Ave.}(C).$$

That is, A’s nephew politely ‘speaks up’ to A’s friend, and A’s friend in turn, as expected, “speaks down” to A’s nephew (see Figure 1: 2a and 2b or 4b1) and 4b2). Specifically, B produces all 15 *Kasar* terms chosen instead of *Lemes respect*. A’s nephew chooses only *Lemes* terms (see Appendix 4).

Upon closer examination, another phenomenon besides status or age differences becomes evident. The following excerpt from text 48 illustrates this. **Lemes** words are in bold roman, **Kasar** words in bold italics.

- 1 C: *Aya naon ieu, Ibi, ribut-ribut kieu?*
there.is what this auntie noisy-noisy like.this
What's up here, Auntie, so noisy like this?
- 2 B: *Nyaéta ibi téh hayang di-gawé yeuh.*
that.is auntie FOC wants PASS-work y'know
Well, I want to work, y'know.
- 3 A: *Heueuh, cenah hayang di-gawé nu gedé buruh-na, aya, Dir?*⁶
unh-hunh says wants PASS-work which large pay-DEF is.there Dir
Ya, says she wants work with a large salary, is there work like that, Chaidir?
- 4 C: *Badé moal, Ibi?*
want.to will.not auntie
You do want work, do you, Auntie?

A's nephew, C, answers A's (C's auntie) and B's (A's friend) questions in *Lemes* style. Occasionally, B uses a *Lemes* term, but it is always the same term just used by C.

- 8 C: *Ieu badé moal?*
this want.to will.not
Do you want this or not?
- 9 B: *Nya badé waé ari taeun mah.*
y'know want.to indeed with.regard.to that FOC
Y'know I do indeed want what you were talking about.
- 10 A: *Enya. Siti ambéh tong ngalamun waé, kaluman némpo-na.*
y'know Siti so.that don't daydream just uneasy see-it
Y'know. Siti, don't just daydream. I'm uncomfortable seeing it.
- 11 B: *Kesel ngalamun waé.*
disgusted daydream just
I'm disgusted just sitting around daydreaming.
- 12 C: *Atuh énjing wé ku abdi.*
indeed tomorrow just by me
Well, I'll find a job for you tomorrow.
- 13 B: *Ah entong énjing mending gé ayeuna waé.*
ah don't tomorrow better also now just
Ah, don't talk of tomorrow, it'd be better to do something just now.

.....

⁶ Speaker A uses *Panengah* style to refer to B's activities, thus symbolising their close relationship (see Figure 2, Type IV).

- 17 A. **Gedé** gaji-na?
Large salary-DEF
- 18 C: **Dua rébu sa-dinten mah.**
two thousand one-day FOC
2,000 rupiahs per day.
- 19 A. **Ah. Lumayan atuh.**
ah not.bad indeed
Ah, not bad at all.
- 20 B. **Dua rébu sa-dinten?! Euleuh untung atuh nya.**
two thousand one-day oh.my profitable indeed y'know
2,000 per day?! Oh my! Good fortune indeed, you know.

Lemes terms (**badé**, line 9, **énjing**, line 13, and **dinten**, line 20) in B's utterances are exceptions to the Satjadibrata's model. But they are motivated, and this motivation can be stated as follows:

I. When repeating a co-speaker's utterance, as for instance, in a clarifying or challenging speech act, repeat the exact words. Do not change speaking style, regardless of social relationship.

As the conversation proceeds, C and B conform to stated norms: that is, B uses *Kasar* down to C; C uses *Lemes* up to B. But, yet again, B uses *Lemes* style (line 23), counter to stated norms, this time to express thanks for C's information.

- 23 B. **Oh, nya atuh keun hatur nuhun.**
oh y'know indeed let.be give thanks
Oh, all right, that's it, thank you.
- 24 C: **Sa-wangsul-na éta téh.**
one-return-NOM that FOC
You're welcome for that.

Neither A's relationship to B, nor B's to C warrants *Lemes* style. The act of thanking someone seems to be somehow intrinsically *Lemes* whereby a speaker would know the following:

II. When expressing 'thanks', use *Lemes* style, regardless of any other style being used.

As for the forms **nuhun** and **hatur nuhun**, Soedradjat feels they are really not so different in politeness. So perhaps these days this so-called exception is not such a great exception after all.

In conclusion, then, for ACQUAINTANCES, excluding counter-examples, the averages for interlocutors B and C are 0% and 100%, respectively; perfectly reciprocal; and conforming perfectly to Satjadibrata's norms.

A similar phenomenon occurs among CO-WORKER dyads. In this case, the average for the same dyad, 44a, is different in Scenes 1 (44a.1) and 2 (44a.2) of the conversation. That is, the relationship between interlocutors is constant, but the average for the dyad differs appreciably from Scene 1 to Scene 2. In short, a change in speech-level use occurs from 44a.1 to 44a.2.

In Text 44, Scene 1, two male CO-WORKERS (C and D), slightly different in age, converse in the office. The conversation opens on the topic of work, then shifts to a personal topic, D's son's problems in school. The style is *Kasar*, expected among *well-acquainted persons of equal status/rank* (see Figures 1–4a). **Lemes** words are in bold roman, *Kasar* words in bold italics.

- 1 C: *Ku naon, Dju, éta mani...mani sararedih* kieu?*
 from what Djunaedi that rather rather very.very.sad like.this
 Why, Dju, very...so very sad like this?
 *sedih 'sad', -arar-AUG.
- 2 D: *Enya⁶, Dir, poé⁶ ieu pararusing* pisan euy.*
 y'know Chadir day this extremely.confused very indeed
 True, Dir, today I'm very disoriented indeed.
 *pusing 'confused', -arar-AUG.
- 3 C: *Éta pagawéan⁶ di-tumpuk baé meureun⁵ nya⁶.*
 that work PASS-piled.up just no.doubt y'know
 That work is just piling up y'know.
- 4 D: *Pagawéan⁶ numpuk! Mana deui?*
 Work piles.up! What next?
- 5 *Di imah⁵ pusing, si⁴ Ahmad teu lulus sakola...*
 at home confused ol' Ahmad not pass school
 At home headache, young Ahmad failed school...
- 6 C: *Si⁴ Ahmad saha?*
 young Ahmad whose
 Who is Ahmad?
- 7 D: *Anak⁵.*
 Child.
- 8 C: *Oh, anak⁴.*
 Oh, child.
- 9 *Si⁴ Ahmad anak⁴ Djunaédi anu kelas tilu SMP téa?*
 young Ahmad child Djunaedi which class 3 jr.hi.school that
 Ahmad, Djunaedi's son in class 3 junior high school, that one?

4 = *Kasar* < *Lemes respect* 5 = *Kasar* < *Lemes humble* 6 = *Kasar* < *Lemes Enteng*

In line 10, C and D begin to use *Lemes* terms: *Lemes Enteng* in line 10; *Lemes respect*, in line 13; *Lemes humble* and *Lemes respect* in lines 14–15; and *Lemes Enteng* in lines 17, 19 and 21.

- 10 D: **Muhun¹.**
 Yes.
- 11 C: *Teu lulus ayeuna nya⁶?*
 no pass now true
 He didn't pass, did he?

- 12 D: *Teu lulus.*
not pass
He didn't pass.
- 13 C: *Kumaha tah budak⁴ téh? Badé³ ngulang maksad-na³?*
how alas kid FOC going.to repeat intention-DEF
Oh dear, how about that kid? Do you want him to repeat the grade?
- 14 D: *Piwarang² abdi² mah ieu ngulang, ngan⁶ anjeunna³*
command my FOC this repeat only he
- 15 *alimeun³.*
educated
I insist that he repeat now, just so he gets an education.
- 16 C: *Tah, atuh kitu mah.*
alas indeed like.that FOC
What a shame, if that's the case.
- 17 *Kieu waé, atanapi¹ di sakola SMA anu soré⁶.*
like.this just or in school sr.hi.school which late.afternoon
Either repeat, or attend a late afternoon high school.
- 18 D: *Swasta nya⁶?*
private/non-government y'mean
A private school, y'mean?
- 19 C: *Muhun¹, swasta. Éta téh nampi¹ kénéh, SMA nu soré⁶.*
yes private that FOC receive yet sr.high which late.afternoon
Yes, private. One is accepting students, a late afternoon high school.
- 20 D: *Swasta mana nu saé¹ téh?*
private which.one REL good FOC
Which private school is good?
- 21 C: *Muhun¹, éta mah swasta nu SMA 'Bur' tah upamina¹.*
yes that FOC private REL sr.high 'Bur' alas for.instance
Yes, that private one, 'Bur' High School, for example.

1 = *Lemes Enteng*

2 = *Lemes humble* [*Lemes pikeun nu ngahormat*]

3 = *Lemes respect* [*Lemes pikeun nu dihormat*]

4 = *Kasar* replaces *Lemes respect*

5 = *Kasar* replaces *Lemes humble*

6 = *Kasar* replaces *Lemes Enteng*

The social relationship remains constant. Why, then, a shift to *Lemes* style beginning in line 9? I find no explanation in Satjadibrata. But a change of topic accompanied by a change of style points to a rather obvious explanation: *Lemes* style is a consensual way of treating sensitive, potentially embarrassing (i.e. to D here) topics in a face-saving manner. In other words, *Lemes* style serves as a 'bad news cushion'. This is not surprising, given that, in Sundanese culture, other people's feelings (*perasaan*) are treated very delicately. Being sensitive to others' feelings is a mark of a refined, cultured (i.e. **lemes**) person. The active principle in C's and D's thinking is, then, the following:

III. When CO-WORKERS (perhaps also FRIENDS AND ACQUAINTANCES) discuss a topic which constitutes a threat to any of their feelings (*perasaan*), they should select *Lemes* style to 'cushion the bad news'.

In Text 44, Scene 2, another CO-WORKER enters, an attractive woman approximately the same age as C and W. A look at an excerpt from Appendix 2 as follows will show that the overall average for the dyad 44a.1 is 0.40, whereas, for 44a.2, it is 0.78. This is almost a twofold increase in Ave.(Dyad) for the same two co-speakers.

TABLE 2: FREQUENCIES AND AVERAGES (DYAD AND INTERLOCUTORS) FOR TEXT 44 (CO-WORKERS) (Excerpted from Appendix 2)

Dyad.	Loca- Ave.			Frequency		Average		Frequency*		Average*	
Scene Sex Age tion (Dyad)				Lemes	Kasar	# 1	# 2	Lemes	Kasar	# 1	# 2
44a.1 m OY o				0.40	7/5 11/7	0.39	0.42	2/3	6/3	0.25	0.50
44a.2 m OY o				0.78	9/3 4/0	0.69	1.00	3/2	1/0	0.75	1.00
44b.2 x OY o				0.67	7/11 3/6	0.70	0.65	3/5	0/2	1.00	0.71
44c.2 x OY o				0.77	11/6 4/1	0.73	0.86	7/1	3/0	0.70	1.00

*without Lemes Enteng and Kasar < Lemes Enteng

It appears that the newcomer's presence in Scene 2 affects the Ave. (Dyad) for 44a, and the averages for the interlocutors C and D. The meaning of this seems to be that the male colleagues are trying to present a good image to their colleague of the opposite sex.

D (male) greets W (female) and they begin chatting about their work.

23 W: **Padamelan**¹ *di-tinggal-keun* *kalah ka ngarobrol*,
work PASS.abandoned-CAUS lose to chatting

24 **pagawéan**⁶ *téh!*
work FOC
Work losing out to chatting! Think of the work!

25 C: *Ieu saur-na*³, *Willa, anak-na*⁴ *teu naék... teu lulus.*
this says-he Willa child-his not go.up not pass
Willa says his child didn't pass...didn't graduate.

28 W: *Bodo meureunan*⁴ *étah mah. Atawa*⁴ *bandel.*
stupid apparently that FOC or headstrong
That one's stupid apparently. Or stubborn.

32 W: *Mémang upami*¹ *pameget*³ *mah, kedah*³ *teras*³.
indeed if male FOC must ahead
It's true, if it's a boy, he must stay in school.

33 C: *Saur-na*³ *ieu, budak-na*⁴ *téh alim-eun*³ *neras-keun*³.
says-he this kid-his FOC natural-ADV go.ahead-CAUS

34 *Saur*² *abdi*² *téh, ieu waé kedah*² *di-pilarian*² *sakola nu*
say I FOC this just must PASS-selected school REL

- 35 *soré⁶* *supados¹* *henteu* *nganggur*.
 late.afternoon so.that not idle
 He says that, as far as his child is concerned, he must of course stay in school.
 As for me, I think we must look for a late afternoon school, so he's occupied.
- 36 D: *SMA* *mana* *nu* *saé¹* *nya⁶*, *Willa?*
 high school which.one REL good y'know Willa
 Which high school is good, hm, Willa?
- 37 W: *Indra atanapi¹* *Bur*.
 Indra or Bur.
- 39 C: *Tah, SMA* *Indra gé* *saé¹* *tah*.
 alas high school I also good alas
 Alas, Indra High School also is good.
- 40 D: *O, saé¹*.
 Oh, fine.
- 41 W: *Nya⁶*, *lumayan, Katolik*.
 y'know not.bad Catholic
 Y'know, it's not bad. It's Catholic.
- 42 C: *Éta budak⁴* *upami¹* *nganggur* *éta téh kirang¹* *saé¹*.
 that kid if unoccupied that FOC less good
 If that kid is not busy, that's not good.
- 43 W: *Muhun¹*, *upami¹* *kedah³* *nganggur* *mah kirang¹* *saé¹* *pisan...*
 yes if must unoccupied FOC less good very
 Yes, if he is forced to be unoccupied, it's not good at all...
- 44 D: *Nganggur nya⁶?*
 Unoccupied y'know?
- 45 W: *...janten langkung¹* *bandel*. *Nya⁶* *kitu* *waé, Djun*.
 and.then more stubborn/naughty y'know like.that just Djun
 ...then he'll become more stubborn/naughty. Isn't that true, Djun?
- 46 D: *Muhun¹*, *saé¹* *di-teras-keun³* *waé, nya⁶?*
 yes good PASS-explain-CAUS just y'know
 Yes, it will be good to just go through with it, don't you think?
- 47 C: *Tah kitu* *waé di-carios-keun³* *ka budak-na⁴* *tah, kedah³*
 alas like.that just PASS-clarify-CAUS to child-DEF alas must
- 48 *sakola soré⁶* *da, sami¹* *kénéh, soré⁶* *sareng¹* *énjing¹*
 school late.afternoon hm same still late.afternoon and morning
- 49 *gé sami¹* *da SMA* *mah ka perguruan tinggi-na mah*
 also same hm hi.school FOC to education high-NOM FOC
- 50 *sami¹* *wé* *di-tampi³* *mah*.
 same just PASS-accepted FOC
 Well it's got to be explained to the kid that he must attend late afternoon

school. It's all the same, late afternoon and day class, the same, in high school or in higher education it's the same problem, getting accepted.

1 = <i>Lemes Enteng</i>	4 = <i>Kasar</i> replaces <i>Lemes respect</i>
2 = <i>Lemes humble</i> [<i>Lemes pikeun nu ngahormat</i>]	5 = <i>Kasar</i> replaces <i>Lemes humble</i>
3 = <i>Lemes respect</i> [<i>Lemes pikeun nu dihormat</i>]	6 = <i>Kasar</i> replaces <i>Lemes Enteng</i>

In summary, references to D's son remain in *Kasar* style, as do *Kasar*<*Lemes Enteng* terms, such as (*e*)*nya* 'yes, true'. The remainder of terms are *Lemes* of various types (see Figure 2).

Thus, in addition to the use of *Lemes* style for 'cushioning bad news' already manifest in dyad 44a.1, where both speakers are males, there is a further factor stirred into the mix which further refines the *Lemes* style in 44a.2. This may be stated as follows:

IV. When male CO-WORKERS (perhaps also FRIENDS and ACQUAINTANCES) are speaking and are joined by a woman, they should select *Lemes* style (or a more elevated degree of *Lemes*) in the woman's presence.

How much of this elevation of style is due to 'cushioning' and how much is due to the woman's presence cannot be precisely determined.

In another example, a similar change of Ave.(Dyad) is evident for dyad 49a from Scene 1 (49a.1) to Scene 2 (49a.2).

TABLE 3: FREQUENCIES AND AVERAGES (DYAD AND INTERLOCUTORS) FOR TEXT 49 (CO-WORKERS) (Excerpted from Appendices 2-4)

Dyad.	Scene	Sex	Age	Status	Loca- tion	Ave. (Dyad)	Frequency		Average		Frequency*		Average*	
							<i>Lemes</i>	<i>Kasar</i>	# 1	# 2	<i>Lemes</i>	<i>Kasar</i>	# 1	# 2
49a.1	x	OY			o	0.70	4/12	5/2	0.44	0.86	4/6	4/0	0.50	1.00
49a.2	x	OY			o	0.50	2/4	5/1	0.29	0.80	1/0	1/0	0.50	---
49c.2	x	OY	d		o	0.76	8/8	1/4	0.89	0.67	5/3	0/2	1.00	0.60
49b2	m	OY	d		o	1.00	4/2	0/0	1.00	1.00	3/1	0/0	1.00	1.00

**Lemes respect/hormat* and *Kasar*<*Lemes respect/horma*
For explanation of abbreviations, see notes to Appendix 1.

In 49a.1, two co-workers, D, a 28-year-old man, and W, a 27-year-old woman, discuss a matter in an office. W complains of not having received her monthly salary. **Lemes** terms are in bold roman; ***Kasar*** terms are in bold italics.

- 1 W: *Dju, kumaha yeuh mani l-al-ieur^{2,*} teu boga⁴ duit⁴.*
Dju how hm rather very-confused not have money
Dju, how about it, hm. I'm rather confused, don't have much money.
**Lieur 'confused', -al-AUG.*
- 2 D: **Muhun¹, Will, tos¹ kaping¹ hiji teu acan² nampi² artos² waé.**
yes Will already date one not yet receive money just
Yes, Will, it's already the first of the month, but we haven't received our salary yet.

- 3 W: *Kumaha atuh majikan téh mani kitu pisan,*
how indeed boss FOC rather like.that very
- 4 *teu di-perhatos-keun² pisan, ka-ter-lalu-an!*
not PASS-pay.attention-CAUS very NOM-too-past-ISER
What a shame the boss's actions are like that, we're totally ignored. It's too much!
- 5 D: *Kumaha nya⁶ majikan téh... majikan téh kitu?*
how y'know boss FOC boss FOC like.that
What about that boss, y'know...what about a boss like that?
- 6 W: *Taros-keun² geura... usul.*
ask-CAUS quickly suggestion
Ask about it soon...I suggest.
- 7 D: *Muhun¹, upami¹... engké siang¹ ieu (heu)nteu hasil engké abdi²*
yes if later day this not succeed later I
- 8 *ka-ditu lah ka majikan, ka dunungan.*
to-there hm to boss to employer
Yes, if...later in the day this is not solved, later I'll go there to the boss, to our employer.
- 9 W: *Kumaha ari⁵ abdi² mah répot² atuh mani kacida⁵ pisan?*
how if I FOC busy indeed rather busy very
What about my being so busy, so very busy?
- 10 D: *Nya⁶ sami³ waé, Will, sareng² abdi².*
y'know same just Will as me
True, it's just the same, Willa, as for me.

1 = *Lemes Enteng*

4 = *Kasar* replaces *Lemes respect*

2 = *Lemes humble* [*Lemes pikeun nu ngahormat*]

5 = *Kasar* replaces *Lemes humble*

3 = *Lemes respect* [*Lemes pikeun nu dihormat*]

6 = *Kasar* replaces *Lemes Enteng*

In Scene 2, the boss joins the conversation (dyads 49a.2, 49b.2 and 49c.2). He asks what they are chatting about.

- 15 W: *Éta, Pa... biasa lah mani tos¹... kosong...*
that sir usual hm rather already empty
That, sir...the usual, it's already...rather...empty.
- 16 C: *Oh, perkawis¹ gaji.*
oh matter salary
Oh, the matter of salary.
- 17 W: *Muhun¹, tos¹ kaping¹ hiji, Bapa mah mani*
yes already date one you.sir FOC rather
- 18 *sa-ngeunah-na⁴ pisan.*
as-comfortable-ADV very
Yes, it's already the first of the month, and you, sir, are very lacksadaisical.

19 C: *Kumaha nya⁶ kersa³ ngantosan³ dua dinten¹ deui?*
 how y'know must wait two days more
 How about it hm, must we wait two more days?

20 W: *Badé³ ka mana kitu?*
 want.to to where like.that
 Where are you off to, if that's the case?

.....

22 C: *Nyaéta bapa téh badé² ka-luar kota ayeuna téh.*
 that.is sir FOC want.to go-out city now FOC
 Well, I'm going out of town now.

.....

24 W: *Bapa mah nyanyabaan⁴ waé ari⁶ urang... lalieur² yeuh.*
 you.sir FOC go.far.away just as.for us confused yah
 You, sir, are going far away and we're very confused y'know.

25 C: *Ieu aya urusan penting.*
 this there.is matter important
 This is an important matter.

26 D: *Ka-luar kota tah dua dinten¹ deui! Ieu tos² ngapngapan yeuh!*
 to-out city alas two day more this already gasping yah
 Going out of town two more days! We're already gasping for air!

27 C: *Tuda bapa téh kedah² ayeuna, ieu téh, berangkat² ka-luar*
 because sir FOC must now this FOC travel to-out

28 *kota téh, margi¹ ieu urusan parusahaan.*
 city FOC because this matter business
 I must now leave to go out of town, because it's a company matter.

29 W: *Asal⁶ ulah bohong wé, Pa, da tos¹... ieu pisan atuh...*
 basis don't lie just, sir 'cause already this very alas

30 *répot² pisan.*
 critical very
 Just don't lie, sir, because already...we are in a very, alas...critical situation.

31 C: *Sing salabar wé heula-nan ayeuna mah.*
 well very.patient just first-ADV now FOC
 You both must just be patient now.

32 D: *Muhun¹ sabar lah.*
 yes patient hmm
 Yes be patient.

33 W: *Muhun¹ lah. Kumaha, Djun?*
 yes hmm how Djun
 Yes, indeed. How, Djun?

- 34 D: *Nya*⁶ kumaha, Will?
y'know how Will
Well how about it, Will?
- 35 W: *Kalahka indit*⁴ *geuningan kapala-na nya*⁶.
rather take.off indeed head-DEF y'know
Rather the boss is taking off, y'know.
- 36 D: *Ieu mah ka-paksa ka loak waé*.
this FOC CAUS-forced to pawn just
- 37 *Ieu mah yeuh ngical*¹ *naon-naon wé lah*.
this FOC yah sell what-what just hmm
As for me, I'll be forced to sell something. I must, alas, pawn something.
- 38 W: *Enya*⁶ *meureunan*⁶ *ka-paksa, hayu ah geus*⁶ *beurang*⁶
y'know assuredly CAUS-force let's ah already midday
- 39 *urang uih*[?] *wé lah*.
we go.home just hmm
True, we're forced to pawn something, all right, let's, ah, do it. It's already midday. Let's go home, hm.

1 = *Lemes Enteng*

4 = *Kasar* replaces *Lemes respect*

2 = *Lemes humble* [*Lemes pikeun nu ngahormat*]

5 = *Kasar* replaces *Lemes humble*

3 = *Lemes respect* [*Lemes pikeun nu dihormat*]

6 = *Kasar* replaces *Lemes Enteng*

For the D-W dyad (49a), the average drops from 0.70 (49a.1) to 0.50 in (49a.2). The boss's presence is clearly a factor in the symbolic social meaning system. Closer inspection reveals the shift in average to be due to W's indignation (lines 17 and 24), by choosing a *Kasar* word, where she would normally select a *Lemes respect* term, according to Satjadibrata.

The boss uses *Lemes enteng* (lines 16, 19 and 28), *Lemes respect* (line 19), and *Lemes humble* (lines 22 and 27). This is unexpected, given that C is W's and D's boss! It appears that C is attempting to deflect W's anger, after she hears of C's impending two-day out-of-town business trip. Careful observation leads us to infer two more 'exceptions' to stated norms, both relating to anger:

V. Use *Kasar* style to express anger in extreme circumstances, even to one's superior.

and,

VI. Use *Lemes* style to deflect anger, regardless of relative statuses of interlocutors.

C leaves and W's anger increases. C and D converse in *Kasar* style in accord with stated norms. When W refers to the boss's going away, according to Satjadibrata she should use *Lemes* style (indicating respect), but uses *Kasar* (line 35: *indit*) instead.

Of course, use of levels to vent anger is tacitly known by native-speakers, whose immediate response is, "Yes, that's right". The point to be made is that this knowledge, readily confirmed by native speakers, does not appear in explicit statements of speech-level norms. Thus, readers or learners with no first-hand experience of Sundanese remain

ignorant of so-called exceptions of the type which Spradley and McCurdy (1975:29) have referred to as “rules for breaking rules”.

iii) FRIENDS AND FAMILY (Refer to Appendices 5–8)

Overall averages for FRIEND and FAMILY member dyads are virtually the same:

$$\text{Ave.}(\text{FRIENDS}) = 0.77 < 0.75 = \text{Ave.}(\text{FAMILY members}).$$

In contrast to STRANGER, ACQUAINTANCE and CO-WORKER dyads, there is a much wider range of averages among FRIENDS and FAMILY member dyads, no doubt because a wider range of emotions, from intimacy to strict decorum, is possible among FRIENDS and FAMILY members.

More interlocutors' averages are balanced for FRIEND dyads than for FAMILY member dyads. This is true for calculations based on **all level variants** (see Appendix 6) or on calculations based only on *Lemes respect/humble* and *Kasar<Lemes respect/humble* terms (see Appendix 7). The status *friend* seems sufficient explanation for this balance. The slight variation between averages for both speakers is due to third person references (i.e. of higher/lower status than speaker and/or addressee).

Still another example to be examined is dyad 58a.1. The overall average for this dyad is located toward the *Lemes* end of the scale of averages for FRIEND dyads:

$$\text{Ave.}(58a.1) = 0.97; \text{Ave.}(\text{Interloc. 1}) = 1.00; \text{Ave.}(\text{Interloc. 2}) = 0.97 \quad [\text{all terms}]$$

$$\text{Ave.}(\text{Interloc. 1}) = 1.00; \text{Ave.}(\text{Interloc. 2}) = 1.00$$

[*Lemes respect/humble* and *Kasar<Lemes respect/humble* terms only]

TABLE 4: FREQUENCIES AND AVERAGES (DYAD AND INTERLOCUTORS) FOR DYAD 58a.1 (FRIENDS) (Excerpted from Appendices 5–8)

					<u>Frequency</u>		<u>Average</u>		<u>Frequency*</u>		<u>Average*</u>	
Dyad.		Loca-		Ave.			Interlocutor				Interlocutor	
Scene	Sex	Age	tion	(Dyad)	<u>Lemes</u>	<u>Kasar</u>	<u># 1</u>	<u># 2</u>	<u>Lemes</u>	<u>Kasar</u>	<u># 1</u>	<u># 2</u>
58a.1	m	OY	i	0.97	16/20	16/21	1.00	0.95	8/5	0/0	1.00	1.00

**Lemes* *respect/humble* and *Kasar* < *Lemes* *respect/humble* terms only

D is C's older friend. Both are males. D wants to know about the yield from C's rice fields this year. C reports it has been a rather good year. D congratulates C.

3 D: Saé¹ nya⁶ ayeuna.
Fine, y'know, now.

4 C: Sumahun¹. Kumaha yeuh sawah Akang anu di Cianjur?
yes how yah rice.field older.brother REL in Cianjur
Yes, older brother, how about your rice field in Cianjur?

5 D: Numawi¹ ieu, Yi, kagung-an² akang mah... rada
and.so this younger.bro. be-owned older.brother FOC rather

- 6 **kirang¹ saé¹ taun ieu téh hasil-na.**
 less good year this FOC harvest-DEF
 And so, younger brother, mine...the harvest was not so good this year.
- 7 C: **Ku-naon?**
 from-what?
 How so?
- 8 D: **Margi¹... keuna hama wereng.**
 because struck disease pest
 Because...there was a plant disease causing insect pest.
- 9 C: **Atuh... kedah³ di... ieu... di-pasihan³ pupuk... sareng¹**
 goodness must be- this PASS-applied fertiliser and
- 10 **di-basmi waé hama wereng.**
 PASS-sprayed just disease pest
 Goodness...it must ah...be fertilised...and sprayed for plant disease.
- 11 D: **Oh, pupuk, pupuk mah parantos². Pupuk sareng¹... éh,**
 oh fertiliser fertiliser FOC already fertiliser and unh
- 12 **parantos² di-semprot. Namung¹... rupa-na¹ teu acan³**
 already PASS-sprayed but appearance-ADV not yet
- 13 **di-pasihan³ sae¹ ku Gusti panginten¹.**
 PASS-given good by Lord probably
 Oh, fertiliser, I've already applied fertiliser. Fertiliser and...ah, and it's already been sprayed. But...apparently not yet blessed by God.
- 14 C: **Sumuhun¹. Ah, éta mah kedah³... ieu wé di-pa-damel³-na.**
 yes ah that FOC must this just PASS-do-it
 Yes. Ah, that must...immediately be done.
- 15 D: **Dupi³... Ayi kumaha tah, dugi¹ ka saé¹ kitu hasil-na?**
 as.for y'er.brother how alas attain to good like.that harvest-the
- 16 **Saratsarat-na kitu téh?**
 requirement.PL-DEF like.that FOC
 As for...you, younger brother, how about it, to get such a good result, how did you do it?
- 17 C: **Ah, teu aya² sarat-sarat abdi² mah... éta waé mung¹**
 ah not there.are requirement.PL my FOC that just but
- 18 **di-pasihan³ pupuk.**
 PASS-given fertiliser
 Ah, it was nothing special...just seeing that the field was fertilised.

.....

1 = Lemes Enteng

2 = Lemes humble [Lemes pikeun nu ngahormat]

3 = Lemes respect [Lemes pikeun nu dihormat] 6 = Kasar replaces Lemes Enteng

Though this is a conversation between male friends, the average is relatively high for this dyad. It appears that again the high average is due to the topic, one which requires *Lemes* style to 'cushion bad news'.

The average for Text 50.1 is mid-range among averages for FRIEND dyads:

Ave. = 0.66; Ave.(Interlocutor 1) = 0.67; Ave.(Interlocutor 2) = 0.64.

TABLE 5: FREQUENCIES AND AVERAGES (DYAD AND INTERLOCUTORS) FOR DYAD 50a.1 (FRIENDS) (Excerpted from Appendices 5–8)

				<u>Frequency</u>		<u>Average</u>		<u>Frequency*</u>		<u>Average*</u>		
Dyad.		Loca- Ave.				Interlocutor				Interlocutor		
Scene	Sex	Age	tion (Dyad)	<u>Lemes</u>	<u>Kasar</u>	<u># 1</u>	<u># 2</u>	<u>Lemes</u>	<u>Kasar</u>	<u># 1</u>	<u># 2</u>	
50a.1	m	OY	p	0.66	12/7	6/4	0.67	0.64	7/1	3/3	0.70	0.25

**Lemes* respect/humble and *Kasar*<*Lemes* respect/humble terms only.

C and D are friends. One day, C comes upon D who looks at his wits end.

- 1 C: **Badé**³ *ka-mana, Dju, luntanglantung kieu?*
want.to to-where Dju strolling.aimlessly like.this
Where to, Dju, wandering around like this?
-
- 4 D: **Puguh**⁶ *ieu, Dir, kamari téh, ka-banjiran euy, imah*⁵ *téh.*
clearly this Dir yesterday FOC CAUS-flood alas home FOC
As you can see, (Chai)dir, yesterday, alas, the house was flooded.
- 5 C: *Ka-banjiran?*
CAUS-flood
Flooded?
- 6 D: **Muhun**¹.
Yes.
- 7 C: *Kumaha ari*⁶ *barudak*⁴?*
how as.to children
How about the children?
**Budah* 'child', -ar-PL.
- 8 D: **Barudak**⁴ *si*⁴... *salamat*⁴... *mung*¹ *parabot waé... seueur*¹ *nu*
children they safe only implement just much which
- 9 *palid tah... anu p-ar-alid*.*
swept.away alas which washed.away-PL
Children...safe...only the household goods...many washed away.
**Palid* 'washed away' -ar-PL.
- 10 C: *Sukur nya*⁶ *barudak*⁴ *salamat*⁴ *Tuh da ieu hujan nya*⁶
praise y'know children safe goodness because this rain y'know
- 11 **teras-teras-an**¹ *ieu tilu dinten*¹ *ieu.*
continue-AUG-ADV this three day this.

Thank goodness, y'know, the children are safe. My goodness this rain, y'know, has been coming constantly for three days.

12/14 D: *Hujan-na ageung¹ pisan sih... ...teras-teras-an¹ deui.*
 rain-DEF large very hmm continue-AUG-ADV more
 The rain is very heavy...on and on.

1 = <i>Lemes Enteng</i>	4 = <i>Kasar</i> replaces <i>Lemes respect</i>
2 = <i>Lemes humble</i> [<i>Lemes pikeun nu ngahormat</i>]	5 = <i>Kasar</i> replaces <i>Lemes humble</i>
3 = <i>Lemes respect</i> [<i>Lemes pikeun nu dihormat</i>]	6 = <i>Kasar</i> replaces <i>Lemes Enteng</i>

In this text, there is a mixture of 'cushioning bad news' and expressing anger.

Finally, the average for dyad 48a.1 is at the *Kasar* end of the average range for FRIENDS:

TABLE 6: FREQUENCIES AND AVERAGES (DYAD AND INTERLOCUTORS)
 FOR DYAD 48a.1 (FRIENDS) (Excerpted from Appendices 5–8)

					<u>Frequency</u>		<u>Average</u>		<u>Frequency*</u>		<u>Average*</u>	
Dyad.		Loca-		Ave.			Interlocutor				Interlocutor	
Scene	Sex	Age	tion	(Dyad)	<u>Lemes</u>	<u>Kasar</u>	<u># 1</u>	<u># 2</u>	<u>Lemes</u>	<u>Kasar</u>	<u># 1</u>	<u># 2</u>
48a.1	f	OY	i	0.09	2/1	11/18	0.15	0.05	0/1	10/5	0.00	0.17
* <i>Lemes</i> respect/humble and <i>Kasar</i> < <i>Lemes</i> respect/humble terms only.												

**Lemes respect/humble* and *Kasar* < *Lemes respect/humble* terms only.

S is W's older friend. Both are women. W inquires about S's daydreaming. S replies.

- 2 S: *Nyaéta ieu téh mikiran⁵ hayang⁵ di-gawé⁵ jiga batur⁴,*
 that.is this FOC thinking want.to PASS-working like others
- 3 *resep isuk-isuk⁶ ngabring, jadi mun⁵ boga⁵ gawé⁵*
 fun morning-AUG walk.together so just have work
- 4 *mah pan meureunan⁵ rada senang kana haté⁵.*
 FOC then maybe rather happy in liver
 Well, I was thinking I'd like to get work like others. It would be fun every morning walking to work in a group. So having work, would make me happy through and through.
- 5 W: *Éh, ari⁶... sugan⁶ téh di-gawé⁴ da ka-tingal-i¹ basa éta*
 hey as.for thought FOC PASS-work hmm view-ADV time that
- 6 *mah sinarieun sok indit⁴.*
 FOC unexpectedly.happening often go
 Hey, I thought you were working, because I saw you recently after going out.
- 7 S: *Nyaéta ieu téh jadi panganggur. Hayang⁵ téh kitu*
 that.is this FOC become unemployed want.to FOC like.that
- 8 *di-mana waé ngababubabu kitu.*
 at-where just be.household.helper like.that
 That is, I've become unemployed. I hope to work anywhere as a household helper like.

- 9 W: *Énya*⁶ *pan di imah*⁴ *gé loba*⁶ *pagawéan*⁶*na*
y'know isn't at home also runs work then
- 10 *pusingpusing teuing?*
upset very
Don't you also have much work at home? Why such a big fuss?
- 11 S: *Laah ari*⁶ *di imah*⁵ *mah da nyaéta* *pagawéan*⁶*-ana* *teu aya*¹
hm as.for at home FOC alas that.is work-DEF not exist
- 12 *buruh-na geura ari*⁶ *di-gawé*⁵ *mah kajeun teuing capé*⁵
income-DEF quick as.for PASS-work FOC no.matter very tired
- 13 *aya buruh.*
there.is income
Well, as for work at home, that's work without any income. If I'm employed, although I'll get tired, there's some income.
- 14 W: *Nya*⁶ *néangan*⁵ *waé atuh... ké geura... pan seueur*¹
y'know looking.for just indeed later quickly not.true many
- 15 *kenal-an ongkoh.*
acquaint-ance individually
Why don't you just look for work...wait a minute...you have many acquaintances.

The mood here is dictated by S's frustration and anger at being unemployed. This is symbolised by the profusion of *Kasar* terms.

Turning to FAMILY members, ranges of averages for FRIEND and FAMILY member dyads are identical. The similarity stops there, though. Patterns of overall averages and interlocutors' averages are different in FAMILY member and FRIEND dyads. That is, the overall average for FRIENDS is 0.77 and for FAMILY members 0.75. Averages for interlocutors, however, are differently patterned. If we group the members of each dyad who have the higher individual average in one group, and the members who have a lower individual average in another, we can derive a group average for each. The result is a pair of ratios: first, 82:69 for FRIENDS, and secondly, 94:74 for FAMILY members.

The disparity between FAMILY member and FRIEND interlocutors appears also if we calculate the sums of the differences between averages for the two interlocutors, and average them. For all types of terms, this average for FAMILY members is 0.24, and for FRIENDS 0.12, a ratio of 2:1 (see Appendix 7). For *Lemes respect/humble* and *Kasar* < *Lemes respect/humble* terms only, the average for FAMILY members is 0.19, while for FRIENDS it is 0.15, a ratio of 1.25:1 (see Appendix 8).

This difference reflects the greater range of variation due to age/generation, direct/indirect descent groups, sex and so forth, which occur within an extended family. FRIEND relationships are generally more uniform than those among FAMILY members.

FAMILY member dyads provide a range of social relationships, across which to observe speech-level use. The same kinds of relationships discussed thus far are reflected in the averages. One apparently undocumented relationship emerges, however, which is symbolised by choice of the words *budak*, *anak* and *murangkalih* when referring to an adolescent. Satjadibrata shows *budak* and *anak* as either *Kasar* or *Lemes humble*,

depending on their use (see Figure 2, Type III). An adult, however, would choose between *budak* or *anak* (=Kasar) and *murangkalih* (=Lemes) for a third person referent. *Lemes humble* applies only to first-person reference (i.e. to oneself).

In Text 26, Scene 1, a mother (N) and father (F) discuss their son's bad report card. The wife uses slightly more *Lemes* terms than the father. For all terms, Ave.(26a.1) = 0.87; Ave.(Father) = 0.67 and the Ave.(Mother) = 0.92. For *Lemes respect/humble* terms only: Ave.(Father) = 0.50 and the Ave.(Mother) = 0.88. This disparity of averages for the interlocutors is no doubt due to the fact that Indonesian wives tend to be younger than their husbands, hence the value of *respect* is appropriate.

Parents' references to their son and his problem reveal an interesting function of speech levels. The father's first reference is a *Kasar* term (*si Kosim*). Using a *Kasar* term could symbolise the referent's perceived lower status or an interlocutor's intimacy with him.

- F. *Mam, kumaha si Kosim teh? Rapot-na awon pisan!*
 ma how ol' Kosim FOC report-DEF bad very
 Mom, what about that Kosim? His report card is very bad!

The mother, in turn, uses the *Lemes* term (*murangkalih*).

- N. *Atuh da murangkalih-na(Lr) males pisan. Sanés kirang warah.*
 goodness hm child-the lazy very not less taught
 Oh my, the boy's very lazy. It's not that nobody's teaching him.

What is the meaning of this *Lemes respect* reference? She may be showing respect to father (=the father's child) by using a *Lemes respect* term. But when the father asks,

- F. *Kinten-kinten naék moal?*
 apparently-AUG go.up will.not
 Probably he'll pass, won't he?

the mother switches and selects a *Kasar* term to refer to the boy and to what he must do to be able to pass to the next grade:

- N. *Ké upami masih tiasa di-leres-keun sareng budak-na(K<Lr)*
 later if still can PASS-correct-CAUS and child-the
kersa(Lr) di-ajar masih tiasa.
 wants.to PASS-educate still can
 Later if it is still possible to straighten out this thing and the kid wants to study, he still can make it.

What is the meaning of this switch in the midst of a dialogue where interlocutors are using *Lemes* style? Relationship alone cannot account for it. The most reasonable explanation is that the mother is shifting to *Kasar* to show disgust at her son's negligence. That is, here is an enactment of Rule IV regarding anger.

Text 45 is the product of an interaction between CO-WORKERS (dyads 45a.1 and 45a.2) involving the same type of reference to one's own child, that is using a *Kasar* term to symbolise anxiety or unhappiness.

S is unhappy. Her successful daughter was chosen to go overseas for further study. A colleague (C) catches S daydreaming and asks her what's wrong. S replies respectfully, but in a way reflecting pride in her daughter.

S: *Ieu pun anak cenah badé ka luar negeri...*
 this my child they.say going.to to out country
 This, my child, they say, will go overseas...

The colleague congratulates S.

C: *Nya saé atuh, ng-iring bingah. Gaduh(Lr) putra(Lr) badé(Lr)*
 y'know good indeed join happy possess child going.to
balajar ka luar negeri mah!
 study to out country FOC
 Y'know, it's really good, I share your happiness. Having a child who's going to study overseas, indeed!
...Janten katingalina éta téh murangkalih(Lr) aya kamajengan...
 and so apparent that FOC child there.is progress
 ...And so apparently that child is getting ahead in life...

But S is torn between pride in her daughter's achievement, anger about the daughter's impending departure, and respect for her colleagues.

S: *Nyaéta ari ka-hoyong mah, ulah cios, kitu...*
 that.is as for NOM-desire FOC don't happen like.that
Gaduh budak téh, teu dua, teu tilu, mung hiji-hiji-na...
 have child FOC not two not three only one-AUG-DEF
 That is, as for my wish, don't happen, like that...I have this kid, not two, not three, only this very one...

Again, S shifts to a *Kasar* term to refer to her daughter.

iv) SPEAKING IN PUBLIC (see Appendix 9)

Finally, speaking in public highlights the relationship of speech levels to speech functions. Situational factors are constant: a restricted number of speakers, one speaker performs at a time, and an audience.

Tasks performed in public vary from: 1) a village head's announcement of an upcoming public speech to be given in a local mosque after evening prayers, 2) a committee head's greeting and welcome to the audience, 3) a request for someone to read from the Qu'ran, 4) an introduction of the speaker at the aforementioned event, 5) the guest speaker's telling about himself, 6) response to the welcome, 7) retelling of humorous experiences in Indonesia, 8) telling of plans for the duration of his stay, 9) delivery of his main address, 10) compliment to those attending, and finally, 11) closing.

The eleven tasks are in rough chronological order, the latter ten constituting a major speech event.

The data recall Satjadibrata's explanation of the use of *Kasar* for scholarly reading material. In general, *Kasar* vocabulary is fuller and broader than *Lemes* vocabulary. In the texts of these 11 tasks, *Kasar* terms are reserved for the visiting speaker's main address about personal health and hygiene. *Kasar* terms appear to serve as a topic indicator for speaking in public, as well. This phenomenon is not only limited to reading material. That is, use of *Kasar* terms means the material is relatively objective and non-personal (Halliday's 'ideational meaning').

Furthermore, our analysis gives a clear picture of how an effective public speaker uses speech levels to prepare the audience for the main topic.

The averages of tasks 1) and 2) show the village head's and committee head's exclusive use of *Lemes* terms (Texts 36.1, 36.2: Average = 1.00) while making announcements. The committee head uses *Lemes* terms exclusively in tasks 3) and 4) (Texts 36.3, 36.4: Average=1.00), no doubt because announcements are mostly concerned with the relationship between announcer and the villagers, and between the announcers, the villagers and the guest speaker, towards whom they show respect. Objective information is subordinate to the social relationships between speech event participants, or as Halliday might say, 'interpersonal meaning' dominates 'ideational meaning'.

Text 36.1 illustrates this style. RT is the *neighbourhood association head*:

Ke-tua rukun tetangga
NOM-old association neighbour,

RW for the citizen's association:

rukun warga
association member.

The text of the recorded announcement is as follows:

RT: *Para Sadérék sa-daya, utami-na warga RW 05, engké wengi*
PL brother one-all primary-ADV member distr. ass'n 05 later evening

badé di-aya-keun pangaosan anu mana pen-ceramah-na,*
going.to PASS-exist-CAUS read.NOM that which AGENT-lecture-DEF

nyaéta Bapa Doktor XYZ ti Amérika.
that.is Mr Doctor XYZ from USA

Friends, especially residents of district 5, this evening a lecture will be presented by Dr XYZ from the USA.

*Aos 'read', *pang...an* NOM.

Pangaosan di-aya-keun ngawit-an tabuh tujuh tilu puluh di
lecture PASS-exist-CAUS begin-ning hour seven three ten at

Masjid Salman.

Mosque Salman

It will take place at 7:30pm at the Salman Mosque.

Sim kuring sa-laku ka-tua RW 05 ngahatur-keun
person I (Lh) as-acting NOM-old distr.ass'n 05 give-CAUS

barébu-rébu nuhun sa-teuacan-na.
NOM.thousand-PL thanks as-not.yet-NOM

I, as head of district 5, express a thousand thanks beforehand.

The guest speaker sprinkles a few *Kasar* terms throughout his self-introduction in task 5) (Text 36.5: Average = 0.97), his response to the welcome in task 6) (Text 37.1: Average = 0.88), sharing of his experiences in task 7) (Texts 38.1, 38.2 and 38.3: Averages = 0.85, 0.68 and 0.89, respectively), telling about his plans in task 8) (Text 39.1: Average = 0.85), telling something good about the area in task 10) (Text 40.1: Average= 0.82), and closing

his speech in task 11) (Text 40.2: Average = 0.78). Mostly, he uses *Lemes humble* terms to refer to himself and his experiences.

This kind of public polite style can be seen in Text 37.1. The guest speaker (TP) opens his speech with the Arabic blessing, "Peace and mercy and the blessing of God be with you". He then continues...

TP: **Sim kuring ngaraos bingah dina wengi ieu, ku aya-na**
 person I (Lh) feel happy in evening this from exist-NOM
pangaosan anu di-aya-keun ku RW 05, anu mana
 read.NOM which PASS-exist-CAUS by distr.ass'n 05 that which
sim kuring atos kénging kahormatan* kanggé masihan wejangan
 person I (Lh) already receive respect.NOM for giving advice
anu mugi-mugi aya mangpaat-na kanggo urang sa-daya-na.
 which maybe there.is valuable-NOM for us one-all-NOM
 I feel happy this evening, by the fact that this pangaosan is being held by
 neighbourhood 5, in that I have the honour of giving advice which hopefully will be
 of value to everyone.
Sim kuring dina kasempetan* ieu badé nyangga-keun hiji
 person I (Lh) in opportune.NOM this want.to propose-CAUS one
poko pedaran ngena-an, atanapi patali jeung kaséhatan*.
 title touch-ing or tied with healthy.NOM
 Given this opportunity, I want to propose a title connected to, or tied to the topic of
 health.
 **Hormat* 'respect', *sempet* 'opportune', *séhat* 'healthy'; *ka...an* NOM.

In Text 37.1, as the speaker approaches his main address (Text 37.2), he uses *Lemes* terms up to the point where he announces the topic of his main address, itself a *Kasar* term. Thereafter, he uses *Kasar* terms. As he is finishing the main address (ideational meaning), he reverts to almost exclusive use of *Lemes* terms (interpersonal meaning). An excerpt from the main address follows.

TP: **Kaséhatan. Para Sadérék sadaya.**
 healthy.NOM PL brother all
 Health. Friends all.
Dina kasempetan ieu sim kuring badé nyangga-keun hiji
 in opportune.NOM this person I (Lh) want.to present-CAUS one
poko pedaran atanapi judul nyaéta Kaséhatan.
 topic or title that.is healthy.NOM
 Having been given this opportunity, I want to present a topic, namely Health.
Lamun ku urang leukeun di-pikir, di-lenyapan, kaséhatan
 suppose by us objectively PASS-thought PASS-pondered healthy.NOM
téh perlu keur sa-kabéh bangsa di dunya.
 FOC necessary for one-all race on earth
 When we reflect a bit, or ponder more deeply, we become aware that health is
 necessary for all peoples around the world.

7. CONCLUSION

This study has confirmed Satjadibrata's model of speech-level use in Sundanese. But it has also broken new ground by documenting seven additions to it which were previously unaccounted for.

These seven additions constitute knowledge Sundanese speakers already have, and which learners ought to acquire. It concerns not only 'knowledge of' the formal *Lemes* and *Kasar* paradigms, but also 'knowledge how to' use these forms according certain fundamental notions of speech events as Hymes has summarised in the well-known acronym S-P-E-A-K-I-N-G.

1) The first addition concerns a situation when one interlocutor *repeats* something another has just said.

I. When repeating a co-speaker's utterance, as for instance, in a clarifying or challenging speech act, repeat the exact words. Do not change speaking style, regardless of social relationship.

We showed an older speaker of higher status repeating a term used by a younger speaker of lower status, that is repeating the *Lemes* term the younger speaker said. According to Satjadibrata, the *Kasar* equivalent would be required by the older speaker's status.

2) The second addition is the speech act of *thanking* someone.

II. When expressing 'thanks', use *Lemes* style, regardless of any other style being used.

An older, higher status speaker may thank a younger, lower status speaker with the *Lemes* '*Hatur nuhun*'. Thanking someone is intrinsically *Lemes*.

3) The third addition concerns the situation when speaking with someone in *Kasar* style, such as a friend or close associate. Switching into *Lemes* style helps *cushion bad news*. It protects that person's feelings.

III. When CO-WORKERS (perhaps also FRIENDS and ACQUAINTANCES) discuss a topic which constitutes a threat to any of their feelings (*perasaan*), they should select *Lemes* style to 'cushion the bad news'.

4) The fourth addition concerns expected *acknowledgement of a female's presence* in mixed-sex conversational groups.

IV. When male CO-WORKERS (perhaps also FRIENDS and ACQUAINTANCES) are speaking and are joined by a woman, they should select *Lemes* style (or a more elevated degree of *Lemes*) in the woman's presence.

Among CO-WORKERS, when a third person having a relationship of CO-WORKER, FRIENDS or FAMILY member enters, that person has the potential to affect speech-level use between the original two speakers, whereas STRANGERS and ACQUAINTANCES do not.

5) The fifth addition concerns use of speech levels to *express strong, negative emotions*. Such displays among FRIENDS and FAMILY members are usually symbolised by use of *Kasar* references to the person on whom these feelings are focused.

V. Use *Kasar* style to express anger in extreme circumstances, even to one's superior.

6) The sixth addition involves use of speech levels as hedges, or qualifiers, or make excuses to *deflect expressions of negative emotion*, as when a boss tries to placate an employee's anger at being paid late.

VI. Use *Lemes* style to deflect anger, regardless of relative statuses of interlocutors.

7) The final addition related to speaking in public. Speakers will usually begin in *Lemes* style to establish a relationship with the audience (Halliday's 'interpersonal metafunction'), then switch to *Kasar* style for the main address (Halliday's 'ideational metafunction'), thus serving as a *topic indicator*.

Finally, STRANGER, ACQUAINTANCE, CO-WORKER, FRIEND and FAMILY member dyads actually encompass a continuum which has been called a range of *degrees of acquaintance*. *Lemes Enteng* terms and *Kasar* replacements for *Lemes Enteng* terms (Figure 2, Type IV) seem to constitute a psychological sensor to a social situation, above and beyond the scale *degrees of acquaintance*.

A table of *Lemes Enteng* term occurrences and their *Kasar* replacements throughout the texts gives confirming evidence of the continuity of the *degrees of acquaintance* scale.

To begin with, when STRANGERS and ACQUAINTANCES speak, very, very few *Kasar* terms are chosen. In most such dialogues, no *Kasar* terms at all occur. In those where they do, speakers appear to be indicating their judgement that the situation is 'not-so-formal'. What is of interest is that, in these types of dialogues, the *Kasar* term used is almost always used in place of a *Lemes Enteng* (i.e. 'neutral polite') term. Even in STRANGER dyads, variability of feelings about social situation is already a factor. *Lemes Enteng* and *Kasar* interchange plays a role throughout the full range of dyad types as follows.

The percentage of *Kasar* terms used in preference to *Lemes Enteng* terms per dyad jumps dramatically moving from STRANGER to ACQUAINTANCE dyads, levels off for CO-WORKER and FRIEND dyads, then again jumps dramatically for FAMILY member dyads (Table 7).

TABLE 7: % KASAR<LEMES ENTENG REPLACEMENTS PER DYAD

	STRANGERS	ACQUAINTANCES	CO-WORKERS	FRIENDS	FAMILY
Number ¹ /	19/	30/	53/	52/	76/
/Dyad ²	/29	/11	/22	/22	/22
Replace ³	0.66	2.73	2.41	2.36	3.45
¹ = Total number of replacements for the respective degree of relationship.					
² = Total number of dyads for the respective degree of relationship.					
³ = Average number of replacements per dyad for the respective degree of relationship.					

The comparison of *Kasar* terms chosen instead of *Lemes Enteng* terms as opposed to those chosen instead of *Lemes respect/hormat* terms also provides an overview of the continuity of the scale of *degrees of acquaintance* (see Table 8).

TABLE 8: RATIO OF *LEMES ENTENG/LEMES RESPECT/HORMAT* TERMS REPLACED BY *KASAR*

Ratio	STRANGER 4.75	ACQUAINTANCE 1.67	CO-WORKER 1.71	FRIEND 0.98	FAMILY 0.78
<i>K<L. Enteng</i>	19/	30/	53/	52/	76/
-----	/	/	/	/	/
<i>K<L. respect/hormat</i>	/4	/18	/31	/53	/97

Moving from STRANGER dyads to FAMILY member dyads, this ratio decreases from almost 5:1 to a ratio of less than 1:1.

Thus, the influence of social relationships (reflected in the choice of Lh, Lr, K<Lh and K<Lr terms) on the differential use of speech levels only really plays a role for ACQUAINTANCE, CO-WORKER, FRIEND and FAMILY member dyads. On the other hand, a general feel for formality/informality of the speech situation (reflected in the choice of LE and K<LE terms) encompasses the entire scale of *degree of relatedness*.

8. DIRECTIONS FOR FUTURE RESEARCH

An expanded understanding of Sundanese speech-level use now enables a deeper understanding of other aspects of linguistic variation, for example full versus reduced forms (i.e. *parantos* : *atos* : *tos* 'already'), which, besides adding to our understanding of the knowledge Sundanese speakers have in communicating with each other, will also elucidate knowledge required for the use of the speech levels by learners of Sundanese.

APPENDIX 1: FREQUENCIES AND AVERAGES FOR DYADS AND INTERLOCUTORS
(STRANGERS)

Dyad.	Scene	Sex	Rel.	Sta-	Loc	FREQUENCY				Ave.	Interlocutor 1				Interlocutor 2				STR-> ->ACQ?*	Lh&r	K<- Lh&r			
						LP*	L*	K*	Tot.		Frequency				Frequency									
											LP*	L*	K*	Tot.	Ave.	LP*	L*	K*				Tot.	Ave.	
STR	1.3	M	AD	D	I	I	5		6	1.17										No				
STR	1.2	E	OY	D	I	I	11		12	1.08										No				
STR	1.4	M	SAD	S	I	I	16		17	1.06										No				
STR	3.1	M	AD	P	I	I	18		19	1.05										Yes				
STR	28c.2	X	OY	H			20	0	20	1.00	7	0	7	1.00		13	0	13		I	Yes	4 7	0 0	
STR	1.1	M	OY	D	P		21		21	1.00										No				
STR	2.1	E	AD	P			17		17	1.00										Yes				
STR	4.4	E	AD	P			12		12	1.00										Yes				
STR	21a.1	M	OY	P			17		17	1.00										No				
STR	22a.1	X	OY	D	I		13		13	1.00										Yes				
STR	22a.2	X	OY	D	I		18		18	1.00	10		10	1.00		8		8		1.00	Yes			
STR	23a.1	F	OY	D	I		18		18	1.00				1.00						1.00	No			
STR	24a.1	X	OY		I		23	0	23	1.00	8	0	8	1.00		15	0	15		1.00	No	5 8	0 0	
STR	30.2	X	OY		P		35		35	1.00	8		8	1.00		27		27		1.00	No			
STR	52.2	F	SAD	S	O		13		13	1.00	5		5	1.00		8		8		1.00	No			
STR	53.2	X	SAD	S	P		9		9	1.00	7		7	1.00		2		2		1.00	Yes			
STR	24a.2	X	OY		I		45	1	46	.98	29	0	29	1.00		16	1	17		.94	No	14 13	0 0	
STR	23.2	X	OY	D	I		31	1	32	.97	17	1	18	.94		14	0	14		1.00	Yes>Kasar			
STR	5.1	E	AD		P		24	1	25	.96	20	1	21	.95		4	0	4		1.00	Yes>Kasar			
STR	35.2	X	OY		P		21	1	22	.95	9	0	9	1.00		12	1	13		.92	No			
STR	56.2	X	OY		P	I	18	2	21	.95	10	0	10	1.00		1	8	11		.91	No			
STR	54a.1	F	SAD	P	I	33	3	37	.95	1	13	3	17	.88		20	0	20		1.00	Yes>Kasar			
STR	57.2	X	OY		O		16	1	17	.94	4	1	5	.80		12	0	12		1.00	Yes>Kasar			
STR	21.2	X	SAD		P		14	1	15	.93	3	0	3	1.00		11	1	12		.92	No			
STR	34.2	M	OY		P		11	1	12	.92	3	0	3	1.00		8	1	9		.89	Yes>Kasar			
STR	56.2	X	OY		P	I	19	3	23	.91	9	1	10	.90		1	10	2	13		.92	No		
STR	52.2	X	SAD	S	O		30	3	33	.91	17	2	19	.89		13	1	14		.93	No			
STR	27b.2	F	OY	D	H		17	2	19	.89	15	2	17	.88		2	0	2		1.00	Yes>Kasar	11 0	2 0	
STR	29c.2	X	OY		H		8	1	9	.89	5	1	6	.83		3	0	3		1.00	Yes>Kasar	3 2	0 0	
STR	25a.1	M	OY	D	S		8	2	10	.80	1	0	1	1.00		7	2	9		.78	No			
COLUMN TOTALS =						7	561	23	591															
Sex: M = both spkrs are male						Ave.(STRs) =																		
F = both spkrs are female																								
X = spkrs include both male & female																								
E = spkrs could be either																								
** Do STRs introduce themselves, thus becoming ACQs?																								
						STATUS:																		
						D = different						{ LP = Lemes Pisan												
						S = same						{ L = Lemes												
												{ K = Kasar												
												LOCATION:												
												I = near home												
												P = public place												
												O = office												
												H = at home												
												S = at store/shop												

Dyad Scene numbers identify the text number and dyad (in texts 21-60 only) in the text files. For example, 1.3 indicates text 1, version 3, while 3.1 indicates text 3, version 1, etc.; 21a.1 indicates text 21, dyad a, scene 1 of the text; 29b.2 indicates text 29, dyad b (the second dyad), scene 2 of the text, etc.

APPENDIX 2: FREQUENCIES AND AVERAGES FOR DYADS AND INTERLOCUTORS
(ACQUAINTANCES AND CO-WORKERS)

	Dyad.					Sta- tus	Loc- ate	FREQUENCIES					Ave. (Dyads)	Interlocutor 1				Interlocutor 2				
		Scene	Sex	Age	LP			Lem	Pan	Kas	Total	Frequ.				Frequ.						
												L		K	Tot.	Ave.	L	Pn	K	Tot.	Ave.	
ACQ	22.2	F	OY		I	I	15				16	1.06										
ACQ	4.1	E	AD		P		19				19	1.00										
ACQ	29.2	X	OY		H		14				14	1.00										
ACQ	51.2	X	OY		I		14				14	1.00										
ACQ	34.2	X	OY		P		61		2	63	.97	.39	1	40	.98	.22		1	23	.96		
ACQ	51a.1	X	OY		I		35		2	37	.95	.11	1	12	.92	.24		1	25	.96		
ACQ	59c.2	X	AC	D	H		30		2	32	.94	.19	2	21	.90	.11		0	11	1.00		
ACQ	53a.1	F	SAD	S	P		28		3	31	.90	.22	3	25	.88	.06		0	6	1.00		
ACQ	50.2	X	OY		P		36	1	6	43	.84	.10	1	11	.91	.26	1	5	32	.83		
ACQ	60.2	F	SAD		H		28		10	38	.75	.17	8	25	.68	.11		2	13	.85		
ACQ	48.2	X	OY		I		7		27	34	.21	.2	27	29	.07	.5		0	5	1.00		
CO	32a.1	M	OY	S	SK		28			28	1.00											
CO	35a.1	M	OY		P		11			11	1.00											
CO	23.2	X	SAD		I		11			11	1.00											
CO	33.2	X	SAD	S	P		17			17	1.00											
CO	45.2	X	OY	S	O		11			11	1.00											
CO	49b.2	M	OY	D	O		6		0	6	1.00	.4	0	4	1.00	.2		0	2	1.00		
CO	31.2	X	OY	S	O		45		1	46	.98	.18	1	19	.95	.27		0	27	1.00		
CO	33a.1	F	OY	S	P		30		2	32	.94	.18	2	20	.90	.12		0	12	1.00		
CO	34a.1	X	OY	S	P		28		2	30	.93	.18	0	18	1.00	.10		2	12	.83		
CO	34.2	X	OY	S	P		11		1	12	.92	.3	0	3	1.00	.8		1	9	.89		
CO	28b.2	M	OY	S	H		49		4	53	.92	.22	2	24	.92	.27		2	29	.93		
CO	45.2	X	OY		O		68		7	75	.91	.34	4	38	.89	.34		3	37	.92		
CO	57a.1	X	SAD		O		44		6	50	.88	.27	6	33	.82	.17		0	17	1.00		
CO	31a.1	X	OY	S	O		32		5	37	.86	.15	4	19	.79	.17		1	18	.94		
CO	45a.1	F	OY		O		29		5	34	.85	.18	2	20	.90	.11		3	14	.79		
CO	32.2	M	OY		SK		18		4	22	.82	.11	3	14	.79	.7		1	8	.88		
CO	60a.1	X	OY		H		18		5	23	.78	.11	2	13	.85	.7		3	10	.70		
CO	44a.2	M	OY		O		14		4	18	.78	.9	4	13	.69	.5		0	5	1.00		
CO	44c.2	X	OY		O		17		5	22	.77	.11	4	15	.73	.6		1	7	.86		
CO	49c.2	X	OY	D	O		16		5	21	.76	.8	1	9	.89	.8		4	12	.67		
CO	49a.1	X	OY		O		16		7	23	.70	.4	5	9	.44	.12		2	14	.86		
CO	44b.2	X	OY		O		18		9	27	.67	.7	3	10	.70	.11		6	17	.65		
CO	49a.2	X	OY		O		6		6	12	.50	.2	5	7	.29	.4		1	5	.80		
CO	44a.1	M	OY		O		12		18	30	.40	.7	11	18	.39	.5		7	12	.42		
Column Totals =							1	842	1	148	992	.85										
Ave.(ACQUAINTANCES) =							1	287	1	52	341	.85										
Ave.(CO-WORKERS) =							0	555	0	96	651	.85										

NOTE: See bottom of Appendix 1 for coding conventions.

APPENDIX 3: KASAR<LEMES RESPECT/HUMBLE AND KASAR<LEMES ENTENG
(ACQUAINTANCES AND CO-WORKERS)

Dyad.	Scene	Sex	Age	Sta	Loc	Interlocutor 1				Interlocutor 2				— TOTAL —				Ave.																				
						Kas	Kas	Ave.	Ave.	Kas	Kas	Ave.	Ave.	Kas	Kas	Ave.	Ave.	Ave.	Interlocutors																			
						<h&r	<LEnt	<h&r	<LEnt	<h&r	<LEnt	<h&r	<LEnt	<h&r	<LEnt	<h&r	<LEnt	<h&r	<LEnt	#1	#2																	
ACQ	22.2	F	OY		I																										1.06							
ACQ	4.1	E	AD		P																										1.00							
ACQ	29.2	X	OY		H																										1.00							
ACQ	51.2	X	OY		I																										1.00							
CO	32a.1	M	OY	S	SK																										1.00							
CO	35a.1	M	OY		P																										1.00							
CO	23.2	X	SAD		I																										1.00							
CO	33.2	X	SAD	S	P																										1.00							
CO	45c.2	X	OY	S	O																										1.00							
CO	49b.2	M	OY	D	O	0	0	—	—	0	0	—	—	0	0	—	—	0	0	—	—	0	0	—	—						1.00							
CO	31.2	X	OY	S	O	0	1	.00	1.00	0	0	—	—	0	1	.00	1.00	0	1	.00	1.00	0	1	.00	1.00						.98	.95	1.00					
ACQ	34.2	X	OY		P	0	1	.00	1.00	0	1	.00	1.00	0	2	.00	1.00	0	2	.00	1.00	0	2	.00	1.00						.97	.98	.96					
ACQ	51a.1	X	OY		I	1	0	1.00	.00	0	1	.00	1.00	1	1	.50	.50	1	1	.50	.50	1	1	.50	.50						.95	.92	.96					
CO	33a.1	F	OY	S	P	0	2	.00	1.00	0	0	—	—	0	2	.00	1.00	0	2	.00	1.00	0	2	.00	1.00						.94	.90	1.00					
ACQ	59c.2	X	AC	D	H	0	2	.00	1.00	0	0	—	—	0	2	.00	1.00	0	2	.00	1.00	0	2	.00	1.00						.94	.90	1.00					
CO	34a.1	X	OY	S	P	0	0	—	—	0	2	.00	1.00	0	2	.00	1.00	0	2	.00	1.00	0	2	.00	1.00						.93	1.00	.83					
CO	34.2	X	OY	S	P	0	0	—	—	0	1	.00	1.00	0	1	.00	1.00	0	1	.00	1.00	0	1	.00	1.00						.92	1.00	.89					
CO	45b.2	X	OY		O	3	1	.75	.25	2	1	.67	.33	5	2	.71	.29	5	2	.71	.29	5	2	.71	.29						.92	.89	.92					
CO	28b.2	M	OY	S	H	0	2	.00	1.00	0	2	.00	1.00	0	4	.00	1.00	0	4	.00	1.00	0	4	.00	1.00						.92	.92	.93					
ACQ	53a.1	F	SAD	S	P	0	3	.00	1.00	0	0	—	—	0	3	.00	1.00	0	3	.00	1.00	0	3	.00	1.00						.90	.88	1.00					
CO	57a.1	X	SAD		O	1	5	.17	.83	0	0	—	—	1	5	.17	.83	1	5	.17	.83	1	5	.17	.83						.88	.82	1.00					
CO	31a.1	X	OY	S	O	1	3	.25	.75	0	1	.00	1.00	1	4	.20	.80	1	4	.20	.80	1	4	.20	.80						.86	.79	.94					
CO	45a.1	F	OY		O	2	0	1.00	.00	2	1	.67	.33	4	1	.80	.20	4	1	.80	.20	4	1	.80	.20						.85	.90	.79					
CO	32.2	M	OY		SK	1	2	.33	.67	0	1	.00	1.00	1	3	.25	.75	1	3	.25	.75	1	3	.25	.75						.84	.79	.88					
ACQ	50.2	X	OY		P	1	0	1.00	.00	2	3	.40	.60	3	3	.50	.50	3	3	.50	.50	3	3	.50	.50						.84	.92	.83					
CO	60a.1	X	OY		H	0	2	.00	1.00	0	3	.00	1.00	0	5	.00	1.00	0	5	.00	1.00	0	5	.00	1.00						.78	.85	.70					
CO	44a.2	M	OY		O	1	3	.25	.75	0	0	—	—	1	3	.25	.75	1	3	.25	.75	1	3	.25	.75						.78	.69	1.00					
CO	44c.2	X	OY		O	3	1	.75	.25	0	1	.00	1.00	3	2	.60	.40	3	2	.60	.40	3	2	.60	.40						.77	.73	.86					
CO	49c.2	X	OY	D	O	0	1	.00	1.00	2	2	.50	.50	2	5	.29	.71	2	5	.29	.71	2	5	.29	.71						.76	.89	.67					
ACQ	60.2	F	SAD		H	1	7	.13	.88	0	2	.00	1.00	1	9	.10	.90	1	9	.10	.90	1	9	.10	.90						.74	.68	.85					
CO	49a.1	X	OY		O	4	1	.80	.20	0	2	.00	1.00	4	3	.57	.43	4	3	.57	.43	4	3	.57	.43						.70	.44	.86					
CO	44b.2	X	OY		O	0	3	.00	1.00	2	4	.33	.67	2	7	.22	.78	2	7	.22	.78	2	7	.22	.78						.67	.70	.65					
CO	49a.2	X	OY		O	1	4	.20	.80	0	1	.00	1.00	1	5	.17	.83	1	5	.17	.83	1	5	.17	.83						.50	.29	.80					
CO	44a.1	M	OY		O	6	5	.55	.45	3	4	.43	.57	9	9	.50	.50	9	9	.50	.50	9	9	.50	.50						.40	.39	.42					
ACQ	48.2	X	OY		I	15	12	.56	.44	0	0	—	—	15	12	.56	.44	15	12	.56	.44	15	12	.56	.44						.21	.07	1.00					
N = 41						61	N = 13						33	N = 54						96																		

APPENDIX 4: KASAR<LEMES RESPECT/HUMBLE AND KASAR<LEMES ENTENG
(ACQUAINTANCES AND CO-WORKERS)

	Dyad.	Scene	Sex	Rel. Age	Sta	Loc	Interloc.1		Interloc.2		--REFERENCE--				
							K <Lh&r	K <LEnt	K <Lh&r	K <LEnt	Int.1-	Int.2	Ave.	Average	
											K < Lh&r	(Dyad)	Int.1	Int.2	
CO	49b.2	M	OY	D	O	O	0	0	0	0		1.00	1.00	1.00	
CO	31.2	X	OY	S	O	O	0	1	0	0		.98	.95	1.00	
ACQ	34.2	X	OY		P	O	0	1	0	1		.97	.98	.96	
CO	33a.1	F	OY	S	P	O	0	2	0	0		.94	.90	1.00	
ACQ	59c.2	X	AC	D	H	O	0	2	0	0		.94	.90	1.00	
CO	34a.1	X	OY	S	P	O	0	0	0	2		.93	1.00	.83	
CO	34.2	X	OY	S	P	O	0	0	0	1		.92	1.00	.89	
CO	28b.2	M	OY	S	H	O	0	2	0	2		.92	.92	.93	
ACQ	53a.1	F	SAD	S	P	O	0	3	0	0		.90	.88	1.00	
CO	60a.1	X	OY		H	O	0	2	0	3		.78	.85	.70	
ACQ	51a.1	X	OY		I	I	0	0	1	1		.95	.92	.96	
CO	57a.1	X	SAD		O	I	5	0	0	1		.88	.82	1.00	
CO	31a.1	X	OY	S	O	I	3	0	1	1		.86	.79	.94	
CO	32a.2	M	OY		SK	I	2	0	1	1		.84	.79	.88	
CO	44a.2	M	OY		O	I	3	0	0	1		.78	.69	1.00	
ACQ	60.2	F	SAD		H	I	7	0	2	1		.74	.68	.85	
CO	49a.2	X	OY		O	I	5	0	1	1		.50	.29	.80	
CO	49c.2	X	OY	D	O	0	1	2	2	2		.76	.89	.67	
CO	44b.2	X	OY		O	0	3	2	4	2		.67	.70	.65	
CO	44c.2	X	OY		O	3	1	0	1	3		.77	.73	.86	
CO	49a.1	X	OY		O	4	1	0	2	4		.70	.44	.86	
ACQ	48.2	X	OY		I	15	12	0	0	15		.21	.07	1.00	
CO	45a.1	F	OY		O	2	0	2	1	0		.85	.90	.79	
CO	45.2	X	OY		O	3	1	2	1	1		.92	.89	.92	
ACQ	50.2	X	OY		P	1	0	2	3	1		.84	.92	.83	
CO	44a.1	M	OY		O	6	5	3	4	3		.40	.39	.42	

Krespect/humble :
Both values = 0

Krespect/humble
At least 1 value = 0

PATTERN I

Krespect/humble
Neither value = 0

PATTERN II

NOTE: See bottom of Appendix 1 for coding conventions.

APPENDIX 5: FREQUENCIES AND AVERAGES FOR DYADS AND INTERLOCUTORS (FRIENDS AND FAMILY MEMBERS)

	Dyad.	Scene	Sex	Rel.	Age	Status	Loc-ate	Frequency				Total	Ave.	
								Lem	Pn	Kas				
FAM	30a.1	M	OY				P	13		0	13	1.00		
FAM	43b.2	X	OY				H	17		0	17	1.00		
FAM	41c.2	M	OY		D		H	16		0	16	1.00		
FR	42a.2	X	SAD				H	14		0	14	1.00		
FR	32c.2	X	OY		S	SK		12		0	12	1.00		
FAM	29a.2	F	OY				H	13		0	13	1.00		
FR	27c.2	X	YAD		S		H	11		0	11	1.00		
FR	43a.2	X	SAD				H	37		1	38	.97		
FR	58a.1	M	OY				I	36		1	37	.97		
FR	56a.1	F	SAD				P	23		1	24	.96		
FAM	59b.2	X	AC		D		H	47		3	50	.94		
FR	42a.1	X	SAD				H	14		1	15	.93		
FR	42c.2	F	OY				H	84		6	90	.93		
FR	53b.2	X	SAD		S		P	13		1	14	.93		
FAM	29a.1	F	OY				H	25		2	27	.93		
FAM	31b.2	F	OY				O	53		5	58	.91		
FAM	55a.1	X	OY				H	21		2	23	.91		
FAM	28a.1	X	OY		S		H	31		3	34	.91		
FAM	60b.2	X	OY				H	10		1	11	.91		
FR	54b.2	X	OY				P	20		2	22	.91		
FAM	57b.2	F	OY				O	20		2	22	.91		
FR	24b.2	X	OY		S		I	9		1	10	.90		
FR	58c.2	X	OY				I	35		4	39	.90		
FR	32b.2	X	SAD		S	SK		15		2	17	.88		
FAM	26c.2	X	OY		D		I	35		5	40	.88		
FAM	26a.1	X	SAD		S		I	13		2	15	.87		
FAM	51b.2	X	OY		D		I	66		12	78	.85		
FR	59a.1	F	CC				H	17		4	21	.81		
FR	59a.2	F	CC				H	8		2	10	.80		
FAM	55a.2	X	OY				H	10		3	13	.77		
FAM	46c.2	X	OY		D		H	19		8	27	.70		
FR	47a.2	M	OY				P	7		3	10	.70		
FAM	27a.1	X	AC		D		H	13		6	19	.68		
FR	50a.1	M	OY				P	19		10	29	.66		
FR	47a.1	M	OY				P	22		1	23	.64		
FAM	41b.2	X	OY		D		H	25		14	39	.64		
FR	47b.2	X	OY				P	8		5	13	.62		
FAM	41a.1	X	OY		D		H	19		12	31	.61		
FR	50a.2	M	OY				P	6		7	13	.46		
FR	47c.2	X	OY				P	15		20	35	.43		
FAM	46a.1	F	OY		D		H	7	3	29	39	.18		
FAM	46b.2	X	OY		D		H	3		14	17	.18		
FAM	46a.2	F	OY		D		H	7		36	43	.16		
FR	48a.2	F	OY				I	3		21	24	.13		
FR	48a.1	F	OY				I	3		29	32	.09		
TOTALS: (FR & FAM) =													.76	
	Interlocutor 1				Interlocutor 2									
	Frequency				Frequency									
	Lem	Pn	Kas	Tot.	Lem	Pn	Kas	Tot.	Lem	Pn	Kas	Tot.	Ave.	
11	0	11	1.00	2	0	2	1.00	bro23-bro25						
11	0	11	1.00	6	0	6	1.00	cosF45-cosM50						
12	0	12	1.00	4	0	4	1.00	nephew-uncle						
10	0	10	1.00	4	0	4	1.00							
5	0	5	1.00	7	0	7	1.00							
11	0	11	1.00	2	0	2	1.00	cosF30-cosF60						
3	0	3	1.00	8	0	8	1.00							
14	1	15	.93	23	0	23	1.00							
16	0	16	1.00	20	1	21	.95							
10	1	11	.91	13	0	13	1.00							
20	0	20	1.00	27	3	30	.90	daughter-father						
4	0	4	1.00	10	1	11	.91							
37	4	41	.90	47	2	49	.96							
11	0	11	1.00	2	1	3	.67							
16	0	16	1.00	9	2	11	.82	cosF30-cosF60						
24	4	28	.86	29	1	30	.97	sis35-sis25						
6	2	8	.75	15	0	15	1.00	wife-husband						
15	0	15	1.00	16	3	19	.84	bro30-sis25						
5	0	5	1.00	5	1	6	.83	bro28-sis24						
15	1	16	.94	5	1	6	.83							
2	1	3	.67	18	1	19	.95	sis28-sis17						
6	0	6	1.00	3	1	4	.75							
24	2	26	.92	11	2	13	.85							
8	1	9	.89	7	1	8	.88							
15	2	17	.88	20	3	23	.87	sis-law-/bro-law						
2	1	3	.67	11	1	12	.92	husband-wife						
47	5	52	.90	19	7	26	.73	aunt-nephew						
8	1	9	.89	9	3	12	.75							
4	1	5	.80	4	1	5	.80							
8	3	11	.73	2	0	2	1.00	wife-husband						
13	6	19	.68	6	2	8	.75	sis23-bro28						
4	1	5	.80	3	2	5	.60							
5	5	10	.50	8	1	9	.89	mother-son						
12	6	18	.67	7	4	11	.64							
3	0	7	10	.30	19	1	5	25	.78					
18	7	25	.72	14	0	14	1.00	sis45-bro28						
2	2	4	.50	6	3	9	.67							
9	12	21	.43	10	0	10	1.00	mother-son						
2	3	5	.40	4	4	8	.50							
10	6	16	.63	5	14	19	.26							
2	3	28	.33	.11	5	0	1	6	.83	mother-daughter				
1	14	15	.07	2	0	2	1.00	mother-son						
2	31	33	.06	5	5	10	.50	mother-daughter						
1	12	13	.08	2	9	11	.18							
2	11	13	.15	1	18	19	.05							
TOTALS: (FR & FAM) =													.81	

NOTE: See bottom of Appendix 1 for coding conventions.

APPENDIX 6: FRIENDS AND FAMILY MEMBERS [LEMES HUMBLE/RESPECT; KASAR<LEMES HUMBLE/RESPECT]

FAM	Dyad.	Sex	Rel.	Sta	Loc-	---Interlocutor 1---							---Interlocutor 2---						
	Scene		Age	us	ate	Lh	Lr	K<Lh	K<Lr	N	Ave.	Lh	Lr	K<Lh	K<Lr	N	Ave.		
FAM	30a.1	M	OY		P	bro23-bro25	1	0	0	0	1	1.00	0	0	0	0	0	—	
FAM	43b.2	X	OY		H	cosF45-cosM50	1	6	0	0	7	1.00	1	2	0	0	3	1.00	
FAM	41c.2	M	OY	D	H	nephew-uncle	0	3	0	0	3	1.00	0	2	0	0	2	1.00	
FAM	29a.2	F	OY		H	cosF30-cosF60	4	3	0	0	7	1.00	0	1	0	0	1	1.00	
FAM	59b.2	X	AC	D	H	daughter-father	1	6	0	0	7	1.00	5	6	0	1	12	.92	
FAM	29a.1	F	OY		H	cosF30-cosF60	4	4	0	0	8	1.00	1	6	0	0	7	1.00	
FAM	31b.2	F	OY		O	sis35-sis25	6	10	2	0	18	.89	5	18	0	0	23	1.00	
FAM	55a.1	X	OY		H	wife-husband	1	1	0	0	2	1.00	5	0	0	0	5	1.00	
FAM	28a.1	X	OY	S	H	bro30-sis25	3	4	0	0	7	1.00	7	2	0	0	9	1.00	
FAM	60b.2	X	OY		H	bro28-sis24	0	1	0	0	1	1.00	1	1	0	0	2	1.00	
FAM	57b.2	F	OY		O	sis28-sis17	0	0	0	0	0	—	9	3	0	0	12	1.00	
FAM	26c.2	X	OY	D	I	sis-law/bro-law	3	6	0	0	9	1.00	6	8	0	0	14	1.00	
FAM	26a.1	X	SAD	S	I	husband-wife	0	1	0	1	2	.50	4	3	0	1	8	.88	
FAM	51b.2	X	OY	D	I	aunt-nephew	6	23	0	1	30	.97	4	7	0	5	16	.69	
FAM	55a.2	X	OY		H	wife-husband	2	3	0	1	6	.83	0	0	0	0	0	—	
FAM	46c.2	X	OY	D	H	sis23-bro28	3	6	1	0	10	.90	1	2	0	1	4	.75	
FAM	27a.1	X	AC	D	H	mother-son	2	1	1	2	6	.50	1	5	0	1	7	.86	
FAM	41b.2	X	OY	D	H	sis45-bro28	8	2	0	5	15	.67	0	5	0	0	5	1.00	
FAM	41a.1	X	OY	D	H	mother-son	6	1	2	0	9	.78	0	5	0	0	5	1.00	
FAM	46a.1	F	OY	D	H	mother-daughter	1	1	1	5	8	.25	2	2	0	0	4	1.00	
FAM	46b.2	X	OY	D	H	mother-son	0	1	3	4	8	.13	0	0	0	0	0	—	
FAM	46a.2	F	OY	D	H	mother-daughter	0	0	3	14	17	.00	3	2	1	1	7	.71	
TOTALS:(FAM) =						52	83	13	33	181		55	80	1	10	146			
TOTALS:(FR)&(FAM)=						103	136	27	51	317		91	144	17	20	272			

FR	Dyad.	Scene	Sex	Rel.	Sta	Loc- ate	-----Interlocutor 1-----					-----Interlocutor 2-----					
							Lh	Lr	K<Lh	K<Lr	Tot.	Ave.	Lh	Lr	K<Lh	K<Lr	Tot.
	42a.2	X	SAD		H	0	3	0	0	3	1.00	0	2	0	0	2	1.00
	32c.2	X	OY	S	SK	1	2	0	0	3	1.00	2	3	0	0	5	1.00
	27c.2	X	YAD	S	H	1	1	0	0	2	1.00	2	2	0	0	4	1.00
	43a.2	X	SAD		H	3	6	0	0	9	1.00	10	4	0	0	14	1.00
	58a.1	M	OY		I	4	4	0	0	8	1.00	1	4	0	0	5	1.00
	56a.1	F	SAD		P	1	5	0	0	6	1.00	0	6	0	0	6	1.00
	42a.1	X	SAD		H	2	1	0	0	3	1.00	2	4	0	0	6	1.00
	42c.2	F	OY		H	10	10	2	1	23	.87	1	21	0	0	22	1.00
	53b.2	X	SAD	S	P	4	4	0	0	8	1.00	1	0	0	0	1	1.00
	54b.2	X	OY		P	3	4	0	0	7	1.00	1	0	0	0	1	1.00
	24b.2	X	OY	S	I	1	3	0	0	4	1.00	0	0	0	0	0	-----
	58c.2	X	OY		I	6	3	0	0	9	1.00	2	4	0	0	6	1.00
	32b.2	X	SAD	S	SK	0	6	1	0	7	.86	1	2	0	0	3	1.00
	59a.1	F	CC		H	0	0	0	0	0	-----	4	2	0	2	8	.75
	59a.2	F	CC		H	2	0	0	0	2	1.00	1	2	0	0	3	1.00
	47a.2	M	OY		P	1	0	0	0	1	1.00	2	0	0	0	2	1.00
	50a.1	M	OY		P	6	1	0	3	10	.70	0	1	3	0	4	.25
	47a.1	M	OY		P	1	0	0	3	4	.25	6	1	2	0	9	.78
	47b.2	X	OY		P	0	0	0	1	1	.00	0	2	0	0	2	1.00
	50a.2	M	OY		P	2	0	0	2	4	.50	0	0	1	0	1	.00
	47c.2	X	OY		P	3	0	2	0	5	.60	0	3	5	1	9	.33
	48a.2	F	OY		I	0	0	2	5	7	.00	0	0	4	3	7	.00
	48a.1	F	OY		I	0	0	7	3	10	.00	0	1	1	4	6	.17
TOTALS : (FRIENDS) =							51	53	14	18	136	36	64	16	10	126	
TOT:(FRIENDS&FAMILY)=							103	136	27	51	317	91	144	17	20	272	

NOTE: See bottom of Appendix 1 for coding conventions.

APPENDIX 7: DIFFERENCE BETWEEN AVERAGES FOR INTERLOCUTORS (FRIENDS AND FAMILY): ALL LEMES AND KASAR TERMS

	Dyad.		Rel.	Sta	Loc-	Diff.Ave's.					
	Scene	Sex	Age	tus	ate	Interlocu-	Grouped	%	Cumul.	%	
						tors 1 & 2					
FAM	30a.1	M	OY		P	.00					bro23-bro25
FAM	43b.2	X	OY		H	.00					cosF45-cosM50
FAM	41c.2	M	OY	D	H	.00					nephew-uncle
FAM	29a.2	F	OY		H	.00					cosF30-cosF60
FAM	26c.2	X	OY	D	I	.01					sis-law/bro-law
FAM	46c.2	X	OY	D	H	.07					sis23-bro28
FAM	59b.2	X	AC	D	H	.10	0-10=7	32			daughter-father
FAM	31b.2	F	OY		O	.11					sis35-sis25
FAM	28a.1	X	OY	S	H	.16					bro30-sis25
FAM	60b.2	X	OY		H	.17					bro28-sis24
FAM	51b.2	X	OY	D	I	.17					aunt-nephew
FAM	29a.1	F	OY		H	.18	11-20=5	23	0-20=12	55	cosF30-cosF60
FAM	55a.1	X	OY		H	.25					wife-husband
FAM	26a.1	X	SAD	S	I	.25					husband-wife
FAM	55a.2	X	OY		H	.27					wife-husband
FAM	57b.2	F	OY		O	.28					sis28-sis17
FAM	41b.2	X	OY	D	H	.28	21-30=5	23	0-30=17	77	sis45-bro28
FAM	27a.1	X	AC	D	H	.39	31-40=1	5	0-40=18	82	mother-son
FAM	46a.2	F	OY	D	H	.44	41-50=1	5	0-50=19	86	mother-daughter
FAM	41a.1	X	OY	D	H	.57	51-60=1	5	0-60=20	91	mother-son
FAM	46a.1	F	OY	D	H	.73	71-80=1	5	0-70=21	95	mother-daughter
FAM	46b.2	X	OY	D	H	.93	91-100=1	5	0-100=22	100	mother-son
Ave.Diff.[Ave.(Int.1)&											
Ave.(Int.2)]=						.24					

	Dyad.		Rel.	Sta	Loc-	Diff.Ave's.					
	Scene	Sex	Age	tus	ate	Interlocu-	Grouped	%	Cumul.	%	
						tors 1 & 2					
FR	42a.2	X	SAD		H	.00					
FR	32c.2	X	OY	S	SK	.00					
FR	27c.2	X	YAD	S	H	.00					
FR	59a.2	F	CC		H	.00					
FR	32b.2	X	SAD	S	SK	.01					
FR	50a.1	M	OY		P	.03					
FR	58a.1	M	OY		I	.05					
FR	42c.2	F	OY		H	.06					
FR	43a.2	X	SAD		H	.07					
FR	58c.2	X	OY		I	.08					
FR	56a.1	F	SAD		P	.09					
FR	42a.1	X	SAD		H	.09					
FR	54b.2	X	OY		P	.10					
FR	50a.2	M	OY		P	.10					
FR	48a.2	F	OY		I	.10					
FR	48a.1	F	OY		I	.10	0-10=16	70			
FR	59a.1	F	CC		H	.14					
FR	47b.2	X	OY		P	.17					
FR	47a.2	M	OY		P	.20	11-20=3	13	0-20=19	83	
FR	24b.2	X	OY	S	I	.25	21-30=1	4	0-30=20	87	
FR	53b.2	X	SAD	S	P	.33					
FR	47c.2	X	OY		P	.36	31-40=2	9	0-40=22	96	
FR	47a.1	M	OY		P	.48	41-50=1	4	0-50=23	100	
Ave.Diff.[Ave.(Int.1)&											
Ave.(Int.2)]=						.12					

NOTE: See bottom of Appendix 1 for coding conventions.

APPENDIX 8: DIFFERENCE BETWEEN AVERAGES FOR INTERLOCUTORS (FRIENDS AND FAMILY): LEMES RESPECT/HUMBLE; KASAR<LEMES RESPECT/HUMBLE ONLY

	Dyad.		Sex	Rel.		St	Loc-ate		Differ. Averages Interlocutors 1&2				
	Scene			Age						Grouped	%	Cumulat.	%
FAM	43b.2	X	OY		H			cosF45-cosM50	0				
FAM	41c.2	M	OY	D	H			nephew-uncle	0				
FAM	29a.2	F	OY		H			cosF30-cosF60	0				
FAM	29a.1	F	OY		H			cosF30-cosF60	0				
FAM	55a.1	X	OY		H			wife-husband	0				
FAM	28a.1	X	OY	S	H			bro30-sis25	0				
FAM	60b.2	X	OY		H			bro28-sis24	0				
FAM	26c.2	X	OY	D	I			sis-law/bro-law	0				
FAM	59b.2	X	AC	D	H			daughter-father	8	0-10=9	50		
FAM	31b.2	F	OY		O			sis35-sis25	11				
FAM	46c.2	X	OY	D	H			sis23-bro28	15	11-20=2	11	0-20=11	61
FAM	41a.1	X	OY	D	H			mother-son	22				
FAM	51b.2	X	OY	D	I			aunt-nephew	28	21-30=2	11	0-30=13	72
FAM	41b.2	X	OY	D	H			sis45-bro28	33				
FAM	27a.1	X	AC	D	H			mother-son	36				
FAM	26a.1	X	SAD	S	I			husband-wife	38	31-40=3	17	0-40=16	89
FAM	46a.2	F	OY	D	H			mother-daughter	71				
FAM	46a.1	F	OY	D	H			mother-daughter	75	71-80=2	11	0-80=18	100
Ave.Diff.Ave's(Interl's1&2)=									19				
									Differ. Averages Interlocutors 1&2				
										Grouped	%	Cumulat.	%
FR	42a.2	X	SAD		H				0				
FR	32c.2	X	OY	S	SK				0				
FR	27c.2	X	YAD	S	H				0				
FR	43a.2	X	SAD		H				0				
FR	58a.1	M	OY		I				0				
FR	56a.1	F	SAD		P				0				
FR	42a.1	X	SAD		H				0				
FR	53b.2	X	SAD	S	P				0				
FR	54b.2	X	OY		P				0				
FR	58c.2	X	OY		I				0				
FR	59a.2	F	CC		H				0				
FR	47a.2	M	OY		P				0				
FR	48a.2	F	OY		I				0	0-10=13	62		
FR	42c.2	F	OY		H				13				
FR	32b.2	X	SAD	S	SK				14				
FR	48a.1	F	OY		I				17	11-20=3	14	0-20=16	76
FR	47c.2	X	OY		P				27	21-30=1	5	0-30=17	81
FR	50a.1	M	OY		P				45				
FR	50a.2	M	OY		P				50	41-50=2	10	0-50=19	90
FR	47a.1	M	OY		P				53	51-60=1	5	0-60=20	95
FR	47b.2	X	OY		P				100	91-100=1	5	0-100=21	100
Ave.Diff.Ave's(Interl's1&2)=									15				
									11				

NOTE: See bottom of Appendix 1 for coding conventions.

APPENDIX 9: SPEAKING BEFORE AN AUDIENCE

	Text.			Pa- ne- ngah	Kasar	Pisan	No.	Ave. (Txt.Pt)	
	Part	Location	Lemes				Terms		SPEAKER
Publ	36.1	Village						1.00	Village Head
Publ	36.2	Mosque						1.00	Committee Head
Publ	36.3	Mosque						1.00	Committee Head
Publ	36.4	Mosque						1.00	Committee Head
Publ	36.5	Mosque	33		1		34	.97	Visiting Speaker
Publ	38.3	Mosque	24		3		27	.89	Visiting Speaker
Publ	37.1	Mosque	14		2		16	.88	Visiting Speaker
Publ	39.1	Mosque	17		3		20	.85	Visiting Speaker
Publ	40.1	Mosque	19	1	4		24	.81	Visiting Speaker
Publ	38.1	Mosque	19	4	4		27	.78	Visiting Speaker
Publ	40.2	Mosque	8			1	9	.78	Visiting Speaker
Publ	38.2	Mosque	13		6		19	.68	Visiting Speaker
Publ	37.2	Mosque	11	1	59		71	.16	Visiting Speaker

COLUMN TOTALS =	158	6	82	1	247	
			Ave.(Public) =		.65	

SEQUENTIAL ORDERING:

Publ	36.1	Village						1.00	Village Head
Publ	36.2	Mosque						1.00	Committee Head
Publ	36.3	Mosque						1.00	Committee Head
Publ	36.4	Mosque						1.00	Committee Head
Publ	36.5	Mosque	33		1		34	.97	Visiting Speaker
Publ	37.1	Mosque	14		2		16	.88	Visiting Speaker
Publ	37.2	Mosque	11	1	59		71	.16	Visiting Speaker
Publ	38.1	Mosque	19	4	4		27	.78	Visiting Speaker
Publ	38.2	Mosque	13		6		19	.68	Visiting Speaker
Publ	38.3	Mosque	24		3		27	.89	Visiting Speaker
Publ	39.1	Mosque	17		3		20	.85	Visiting Speaker
Publ	40.1	Mosque	19	1	4		24	.81	Visiting Speaker
Publ	40.2	Mosque	8			1	9	.78	Visiting Speaker

NOTE: See bottom of Appendix 1 for coding conventions.

REFERENCES

- Ayatrohaedi, 1980, Bentuk hormat dalam bahasa Sunda [The structure of respect in Sundanese]. *Scholarly Publications Series* 3:85–101. Jakarta: Faculty of Arts, University of Indonesia.
- Bakels, Jet, 1989, Mandala-gemeenschappen in West-Java. *Bijdragen tot de Taal-, Land- en Volkenkunde* 145:359–364.
- Djajawiguna, B.R.I., 1978, *Babaran undak-usuk basa* [The description of linguistic politeness levels]. Bandung: Institute for Sundanese Language and Literature.
- Ekadjati, E.S., ed., 1980, *Masyarakat Sunda dan kebudayaannya* [Sundanese society and its culture]. Jakarta: Girimukti Pasaka.
- Eringa, F.S., 1949, Loetoeng kasaroeng [Stories of the phantom ape]. *Verhandelingen van het Koninklijk Instituut voor Taal-, Land-, en Volkenkunde* VIII.
- Errington, J.J., 1985, *Language and social change in Java: linguistic reflexes of modernization in a traditional royal polity*. Monographs in International Studies: Southeast Asia Series No. 65. Ohio University Center for International Studies.
- Geertz, H., 1963, *Indonesian cultures and communities*. New Haven: Human Relations Area Files.
- Harsojo, 1983, Kebudayaan Sunda (Sundanese culture). In R. Koentjaraningrat, ed. *Manusia dan kebudayaan di Indonesia* [Society and culture in Indonesia], 300–326. Jakarta: Djambatan.
- Kats, J. and M. Soeridiradja, 1982, Tingkat bahasa [Language levels]. In J. Kats and M. Soeridiradja *Tata bahasa dan ungkapan Bahasa Sunda* [Grammar and expression in Sundanese], 1–13. Jakarta: Penerbit Djamatan.
- Kern, R.A., 1906, 't Lemes in 't Soendaasch [Politeness in Sundanese]. *Bijdragen tot de Taal-, Land- en Volkenkunde* VII:5.
- Noorduyn, J., 1963, Categories of courtesy in Sundanese. *The Bible Translator* 14:4.186–191.
- Rosidi, A., 1980, Ciri-ciri manusia dan kebudayaan Sunda [Distinguishing characteristics of Sundanese society and culture]. In Ekadjati, ed. 1980, 125–161.
- Satjadibrata, R., 1954, *Kamus basa Sunda* [Sundanese dictionary]. 2nd ed. Jakarta: Ministry of Education and Culture.
- 1956, *Undak-usuk basa Sunda* [Linguistic politeness levels in Sundanese]. 2nd ed. Jakarta: Balai Pustaka
- Soedradjat, T.M., 1986, *Cross-cultural study for Sundanese learners of English*. MA dissertation, Southeast Asian Ministers of Education Organization's Regional Language Centre and the National University of Singapore.
- Spradley, James P. and David W. McCurdy, 1975, *Anthropology: the cultural perspective*. New York: Wiley and Sons.
- Statistical Year Book of Indonesia, 1989–1990. Jakarta: Central Bureau of Statistics.
- Surjadi, A., 1974, *Masyarakat Sunda: budaya dan Problema* [The Sundanese People: Culture and Problems]. Bandung: Penerbit Alumni.
- Wessing, R., 1974, Language levels in Sundanese. *Man* 9:5–22.
- Wirakusumah, R.M. and I.B. Djajawiguna, 1957, *Kandaga tatabahasa* [Treasury of grammar]. Bandung: Ganaco Publishers.

KHMER AND THE THEORY OF MODALITY

FERDINAND DE HAAN

1. INTRODUCTION¹

In this paper, I will investigate the category of modality in Khmer (Cambodian). Apart from brief discussions in grammars, modality in Khmer has not been analysed adequately yet. For reasons that will become clear, it is of value to any theory that tries to deal with modal phenomena across languages. The aim of this paper is twofold; first, to provide a better description of the Khmer modal system, and second, by contrasting the modal system of Khmer with that of English, to introduce a framework especially designed for cross-linguistic comparisons. It will be called here the **continuum** model. Even though English and Khmer share a number of similarities, the areas in which the two languages differ is very striking indeed. It will be argued that a model based on a continuum scale, rather than on discrete categories, offers the best basis for cross-linguistic comparisons.

The categories of modality I will be discussing here are the **epistemic** and **deontic** categories, following Palmer (1986). Epistemic modality deals with the degree of trust the speaker has in the truth of his or her speech utterance, while deontic modality concerns itself with the notions of necessity, obligation and permission. The latter category has often been called **root** modality, but I have chosen to use the term deontic modality here, since root modality deals with much more than just obligation, necessity and permission alone, and in this paper I am focusing on just the true deontic features of the modals.

The paper is divided into three main parts. The first part is a theoretical discussion of modality. The data are mainly from English but some data from Dutch have been included for comparison. This section serves as a foundation for the discussion of the Khmer data in the second part and it also argues for the treatment of modal elements as a system, rather than treating the various modal verbs in isolation. The third part of this paper is a discussion of a special area, namely the interaction of modality and negation. It will be shown that English and Khmer, despite their numerous similarities, differ in their treatment of this interaction.

¹ I am grateful to Bernard Comrie, Joan Bybee, John Hawkins, Roger Woodard and Maria Polinsky for commenting on earlier drafts of this paper. Responsibility for this paper is mine alone.

Unless otherwise stated, the Khmer examples are due to my informant, Mr Van So Chau, a native Khmer speaker from the town of Battambang. The Khmer examples are written in a broad phonological transcription. The following abbreviations are used: INF – infinitive; NEG – negation; PAST – past tense; PERF – perfective aspect; Q – question marker; SG – singular.

2. MODALITY IN ENGLISH

It is safe to say that the modal system of English is the best known and best analysed modal system of any language, a fact which has inspired many books and articles. To mention but a few: Boyd and Thorne (1969), Ehrman (1966), Palmer (1990) and Coates (1983). These books and articles have in common the fact that they try to give an account of the English modal system, with an emphasis on the meanings of the individual modal verbs. None of these analyses is therefore really suited for a cross-linguistic comparison of modal systems. The paper by Boyd and Thorne is devoted to a componential analysis of the modals of their particular (British) English dialect. The modals are defined according to their individual meanings. They are not looked at as a system but rather as individual verbs within a semantic field. Their analysis treats epistemic and deontic modals as different verbs and fails to capture the fact that the same modals are used for both epistemic and deontic modality.² It is therefore unsuited for cross-linguistic analyses.

Better suited for cross-linguistic comparisons is the matrix model. A matrix is a "...multi-dimensional framework, with each dimension indicating some set of related semantic features" (Palmer 1990:17). The appeal of a matrix is that it is a clear representation of the relation between the individual elements. The elements are presented as being part of a system, which gives us a better overview of the individual relations between the verbs.

According to Palmer (1990:18), Twaddell (1960:11) was the first to analyse the English modals with the help of a matrix. His proposal is shown in Chart 1 below:

	Prediction	Possibility, capability, permission	Necessity, requirement, prescription
Absolute, unrestricted	<i>will</i>	<i>can</i>	<i>must</i>
Contingent, inconclusive	<i>shall</i>	<i>may</i>	<i>need</i>
Morally determined		<i>dare</i>	<i>ought</i>

CHART 1: MATRIX ANALYSIS OF ENGLISH MODALS (TWADDELL 1960:11)

Problems with this particular analysis are obvious. It is hard to see the difference between 'unrestricted' and 'inconclusive'. Furthermore, the distinction between *must* and *need* is not clear from the matrix. The difference does not seem to be one of 'unrestricted' versus 'inconclusive'. The same considerations hold for *can* versus *may*. Lastly, the verbs *dare* and *ought* are too different to include them into one 'morally determined' category. In addition, the difference between these two verbs is not one of possibility versus necessity.

The difficulty with a matrix analysis is to decide what kind of semantic features to use. In the analysis cited above, one dimension is concerned with the relative strength of the modals: possibility versus necessity. It is not clear, however, how the notion 'prediction' fits into this scheme. It is no logical part of the opposition weak versus strong modality. For reasons such

² For a more detailed analysis of Boyd and Thorne (1969) see de Haan (1994:41ff.).

as these, the specific analysis of Twaddell's cannot be maintained.³ Of course this does not mean that there is something inherently wrong with matrix analyses. Twaddell's analysis merely shows that it is difficult to come up with the right labels for each category.

Palmer (1990:37) takes a different approach. He combines epistemic, deontic and dynamic modality into a matrix. This is shown in Chart 2:

	Epistemic	Deontic	Dynamic
Possibility	<i>may</i>	<i>may/can</i>	<i>can</i>
Necessity	<i>must</i>	<i>must</i>	
?	<i>will</i>	<i>shall</i>	<i>will</i>

CHART 2: MATRIX ANALYSIS OF ENGLISH MODALS (PALMER 1990:37)

If we compare Palmer's matrix with that of Twaddell, it is easy to see that Palmer's matrix is simpler and more appealing intuitively. His big problem is the category indicated by a question mark. Palmer is forced to include this category because of sentences such as those shown in (1a) and (1b), from Palmer (1990:36,37):

- (1) a. John will be in his office.
 b. You shall have your reward tomorrow.

Palmer is unable to group *will* and *shall* with either possibility or necessity and is therefore forced to set up a third category, but he is at a loss as to what this category should be called. In fact, it seems to me that there is no clear correspondence between epistemic *will* and deontic *shall* and that grouping them into one category is not justified. There is no semantic category close enough to 'possibility' and 'necessity' that modal *will* and *shall* have in common.

Other problems with the matrix analysis are more severe. For instance, what is the place of the verb *should* (or *ought to*) in a matrix? Palmer does not include them in his matrix, and indeed it would be hard to do so, since these verbs occupy a place somewhere between possibility and necessity. It would seem that we could simply add a row of cells to the matrix to incorporate this new category, as seen in Chart 3:

	Epistemic	Deontic	Dynamic
Possibility	<i>may</i>	<i>may/can</i>	<i>can</i>
??	<i>should/ought to</i>	<i>should/ought to</i>	
Necessity	<i>must</i>	<i>must</i>	
?	<i>will</i>	<i>shall</i>	<i>will</i>

CHART 3: MATRIX ANALYSIS OF ENGLISH MODALS, REVISED

This approach has two drawbacks. First, one has to give this new category a meaningful name. More serious is the fact that the matrix becomes too full to give an adequate description

³ In this respect it must be mentioned that in a subsequent edition of his book, published in 1963, Twaddell drops this analysis without any explanation.

of the modal system. The logical simplicity of the original has disappeared, since the vertical dimension of the system is now full of incoherent categories.

The new matrix shown in Chart 3 has created new problems for cross-linguistic comparisons. We now have a matrix with four rows and three columns, but in other languages not all cells might be filled, or some languages might make even finer distinctions within their grammaticised modal system. A matrix model is too rigid to take this into account. Consider the case of Dutch, for instance. Dutch has no verb comparable to English *should*. Instead, various forms of the verb *moeten* 'must' are used. Normally, *moeten* denotes strong modality (epistemic and deontic), but in certain contexts it is weakened. In sentence (2a) below we see the subjunctive of *moeten*, which denotes a weakened obligation. In (2b) the past perfect tense of *moeten* is used to denote (unrealised) obligation in the past.

- (2) a. *Hij zou naar Amsterdam moeten gaan.*
 he shall.3SG.PAST to Amsterdam must.INF go.INF
 He should go to Amsterdam.
- b. *Hij had naar Amsterdam moeten gaan.*
 he have.3SG.PAST to Amsterdam must.INF go.INF
 He should have gone to Amsterdam.

The question is: how do we represent the Dutch system by means of a matrix, such as the one in Chart 3? Basically, we have two options available to us. The first is to disregard the use of *moeten* as weakened obligation and leave the corresponding row in the matrix empty. The justification would be that there is no separate verb available in Dutch that corresponds to English *should*. However, this would create the (utterly false) impression that Dutch lacks a way to express the meaning of weakened obligation.

The second option is to fill the appropriate row of the matrix with the combinations *zou moeten* and *had moeten*, in order to show that those are the corresponding Dutch elements. This would be justified if the matrix were one of meaning instead of being one of forms. If the matrix were designed to deal with meaning, many more elements would need to be added, such as modal adverbs, modal particles and modal paraphrases. Furthermore, we would be missing a generalisation, since the verb *moeten* is used twice, on two different levels. In the matrix model, this can be seen only as an accident, since there is no inherent connection between individual cells.

To sum up this section, a matrix model is more useful for cross-linguistic comparisons than any of the other frameworks we have examined, but it suffers from some drawbacks. The most serious one is its rigidity: discrete cells are used to reflect different verbs and different meanings. There is not always a one-to-one correspondence between **meaning** (represented by cells) and **form** (represented by verbs). To overcome these problems, we should look for a 'matrix without cells'.

In this paper, I will adopt an analysis of modals that will be called here the **continuum** analysis. It has some affinity with so-called scalar models, as developed in Horn (1972) and in use in, for instance, the theory of Functional Grammar (Hengeveld 1987, Siewierska 1991). Like scalar models, the continuum model relies on the fact that there is a gradual difference in intensity among certain modals. For instance in English, the modal *must* is stronger in intensity than the modal *may*. I will argue that the difference between these modals can be best expressed by means of a continuum model, shown in Chart 4:

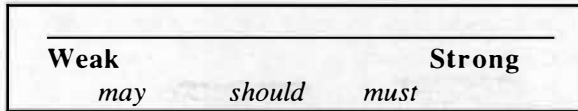


CHART 4: ENGLISH MODALS: CONTINUUM MODEL

In this way, it is easier to represent the relative strength of the modals, since their intensity can be ascertained simply by looking at the relative place on the continuum line. There are no boundaries drawn since in this view the boundaries between the various modal elements are fuzzy. This model is also capable of showing change in the relative place of the modals simply by showing movement of the modal on the continuum line. A diachronic analysis, however, is beyond the scope of this paper.

The question is which elements should get a place on the continuum model. I have chosen to make the model as syntactically homogeneous as possible and put only grammaticised modal elements on the continuum. In this way, I am avoiding all lexical expressions (such as *possibly*, *maybe*, and *certainly*), which would make the continuum model less effective. Items that have a place are verbs (English, Khmer), affixes on the verbs (Tamil), and so on. For a discussion see de Haan (1994:47–49).

The continuum model shows the various modal meanings as parts of a greater whole, not as a hierarchy. In no way is it implied in Chart 4 that strong modality takes precedence over weak modality or that strong modality logically entails weak modality. The above model is simply a linguistic representation, without any implications for logical theory.

This model is well suited for cross-linguistic comparisons, since it provides us with a simple framework on which to place the modal elements. Of course, the continuum model is less suited to account for all the different shades of meaning a given modal may have, but that is not the intention of the model in the first place. It is designed to limit the input (in computational terms) to a few relevant aspects and tries to account for these variations. To account for all different shades of meaning of a given element should not be the task of a cross-linguistic survey of any kind.

For reference, the complete English modal system is shown in Charts 5 and 6 below:⁴

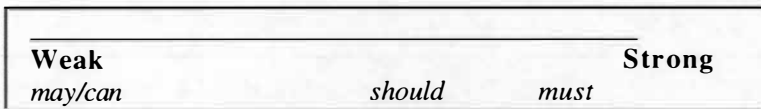


CHART 5: ENGLISH EPISTEMIC MODALS

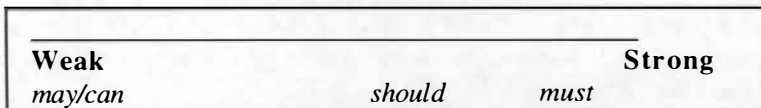


CHART 6: ENGLISH DEONTIC MODALS

⁴ This is a somewhat simplified version of the framework developed in de Haan (1989) for Russian and Chinese and further augmented in de Haan (1992) for Dutch. For the present purposes the version shown in Charts 5 and 6 will suffice to explain the relevant data.

In the next section, this model will be used for the Khmer data, and for a comparison between the English and Khmer systems.

3. THE KHMER DATA

In this section, I will discuss various ways in which modality is expressed in Khmer, though it must be kept in mind that we are concentrating on epistemic and deontic modality. Other modalities will not be discussed. The section is divided into three parts; the first part deals with the verbs, the second part with the adverbs that express modality. The third part is a brief detour into the area of the diachronic status of the Khmer modal system.

For reference, the elements discussed are shown in (3) below:

(3) Khmer modal elements

<i>trəw (taε)</i>	must
<i>trəw kaa</i>	need
<i>kuo (taε)</i>	should
<i>ʔaac</i>	can, be able; may
<i>cam bac (tee)</i>	need
<i>prahæɭ</i>	maybe
<i>pit ciə</i>	must

3.1 DEONTIC MODALITY

As mentioned in the introduction, deontic modality is concerned with the notions of obligation and necessity. In Khmer, these notions are expressed by means of modal verbs. Khmer possesses verbs for the entire spectrum of deontic modal ranges. The verb *trəw (taε)* is used for strong deontic uses; *kuo (taε)* is used to denote the range of meanings of English *should*, while *ʔaac* is used for permission. These verbs can be placed in the continuum model as follows:

Weak		Strong	
<i>ʔaac</i>	<i>kuo (taε)</i>	<i>trəw (taε)</i>	
may	should	must	

CHART 7: KHMER DEONTIC MODALS

The continuum model for the Khmer deontic modal system looks strikingly similar to that of the English system shown in Chart 6 above. Both systems have three verb sets that express the range of deontic notions from weak to strong.⁵ Compare this to the Dutch deontic system shown in Chart 8 below:

⁵ There is a difference between English and Khmer in that English uses two different verbs, *may* and *can*, to express weak deontic modality whereas Khmer only uses one verb, *ʔaac*. Both *may* and *can* are treated here as synonyms because for cross-linguistic purposes both can be said to cover the same range. Of course there are differences between the two English verbs but this is not relevant in the present discussion.

Weak	Strong
<i>mogen/kunnen</i>	<i>moeten</i>
may/can	must

CHART 8: DUTCH DEONTIC MODALS

As discussed in §2, Dutch does not have a separate verb, corresponding to English *should*. Instead, Dutch uses the strong deontic verb *moeten*. We can say that on the continuum scale the verb *moeten* takes up a greater part than *must* does on the corresponding English continuum. Again, a matrix model would not be able to explain this phenomenon while the continuum model accounts neatly for it.

3.1.1 THE VERB *trəw*

We will now turn to a more detailed discussion of the Khmer modal system. The starting point is the strong deontic verb *trəw*. Just like its English counterpart *must* it denotes obligation and necessity, as illustrated by sentence (4) below. Just as in English, the Khmer sentence denotes an obligation or a necessity on the part of the subject to perform a certain action, in this case to kill the fish.

- (4) *Kñom trəw (.tae) səmlap trəy.*
 I must kill fish
 I must kill the fish.

The normal placement of the modal verb is before the main verb, but on occasion it can be found before the subject as well. This is shown in (5):

- (5) *Trəw kñom səmlap trəy.*
 must I kill fish
 I must kill the fish.

The semantic value of sentence (5) is identical to that of (4). The word order in (5) is extremely rare, however, and it may reflect some archaic verbal construction. In the rest of the paper, only the order of Subject – Modal verb – Main verb will be used.

The word order of sentence (5) can also be interpreted as being a question: *Must I kill the fish?* Inversion of subject and verb is normal for questions, although in such cases the sentence-final question particle *rɨi* is normally used. This particle is not obligatory, however. Sentence (5) is then ambiguous between a declarative and an interrogative reading, with the interrogative reading as the preferred interpretation.

A look at the list in (3) shows that *trəw* has two variants, *trəw tae* and *trəw kaa*. Jacob (1968:318) glosses *trəw tae* as ‘absolutely must’ and ‘pre-verbal particle’ as opposed to the glosses ‘must’ and ‘verb’ for *trəw*. The informant I consulted made no distinction between *trəw* and *trəw tae*, however. He used the two forms interchangeably according to stylistic principles, such as smoothness. Occasionally *trəw tae* can be used for emphasis, but the basic difference between the two forms seems to be stylistic in nature. Since both forms are used in the same contexts, I can see no real reason for maintaining Jacob’s distinction between ‘pre-verbal particle’ and ‘verb,’ at least for the dialect that is described here.

The second variant of *trəw* is *trəw kaa*. When used with a main verb it has the meaning of 'to have the duty to'. This is shown in (6) below. When *trəw kaa* is used in this sense, there is a moral obligation or necessity, or an obligation from something (or someone) not present in the situation described. Sentence (6) therefore denotes that the obligation does not come from the speaker himself or herself.

- (6) *Bophaa trəw.kaa səmlap trəy.*
 Bophaa need kill fish
 Bophaa has the duty to kill the fish.

The verb *trəw kaa* has a slightly different distribution than *trəw*, since it can be used with a direct object, while *trəw* cannot. In this respect *trəw kaa* is similar to English *need*. This is shown in (7):

- (7) *Bophaa trəw.kaa trəy.*
 Bophaa need fish
 Bophaa needs a fish.

Khmer has another verb that can be translated by need, (*cam*) *bac*.⁶ It is more restricted to negative sentences. In positive sentences it is synonymous with *trəw*, and can be used interchangeably with it, much in the way that *must* and *need* can be used interchangeably in English positive sentences. In negative sentences, *cam bac* differs from *trəw* with respect to scope of the negation. This is shown in (8) below:

- (8) a. *Kñom mìn trəw səmlap trəy.*
 I NEG must kill fish
 I mustn't/needn't kill the fish.
- b. *Kñom mìn cam.bac səmlap trəy.*
 I NEG need kill fish
 I needn't kill the fish.

A full discussion of negative sentences can be found in §4, but we can observe for now that in sentence (8a) the negation is ambiguous in scope with respect to the modal verb, while the modal verb is in the scope of the negation in sentence (8b). There are some parallels between Khmer and the relative scope of the English verbs *must* and *need* with respect to negation. The situation is more complex, however, due to the fact that the negative particle *mìn* in (8a) can also be placed after the modal verb, with appropriate change in scope. This will be taken up in more detail in §4.

⁶ There are many variants of *cam bac*. The following list was elicited:

positive	negative
<i>bac</i>	<i>mìn bac</i>
<i>cam bac</i>	<i>mìn cam bac</i>
<i>cam bac tee</i>	<i>*mìn cam bac tee</i>

The elements *cam* and *tee*, neither of which can be used by itself, are optional except in the negative, where *tee* is disallowed. The forms with *cam* are more polite and are used when one is speaking to a person of higher rank.

3.1.2 THE VERB *kuo*

The Khmer verb *kuo* has the same denotations as the English *should*. It is used as a weaker form of *trəw*, but otherwise denotes an obligation on the side of the subject to perform a certain action.

- (9) *Bophaa kuo(.tae) səmlap trəy.*
 Bophaa should kill fish
 Bophaa should kill the fish.
- (10) *Bophaa mɨn kuo(.tae) səmlap trəy.*
 Bophaa NEG should kill fish
 Bophaa should not kill the fish.

As is the case with *trəw*, *kuo* can optionally be followed by the particle *tae*. Again, the basic determination seems to be that of style. My informant judged sentence (9) 'smoother' with the inclusion of the particle *tae*, but when a negation is present as in (10), it was judged to be better to omit the particle. Negation will be dealt with further in §4.

3.1.3 THE VERB *ʔaac*

The Khmer verb *ʔaac* has two basic meanings, ability and permission (weak deontic modality). In both its ability sense and its permission sense it functions as an auxiliary verb and cannot be used as a main verb. Thus sentence (11a) is ungrammatical:

- (11)a. **Bophaa ʔaac pyesaa kmae.*
 Bophaa can language Khmer
- b. *Bophaa ʔaac niʔiyey pyesaa kmae.*
 Bophaa can speak language Khmer
 Bophaa can speak Khmer.

Another verb always has to be present to express the nature of the ability, in this case the verb *niʔiyey* 'speak'. Unlike in other languages, such as German, Khmer sentences denoting ability always require a main verb. It makes no difference whether the ability expressed is mental or physical; *ʔaac* + main verb is used in both cases.

The second meaning of *ʔaac* is permission. This is shown in (12):

- (12) *Bophaa ʔaac mɔuc baan.*
 Bophaa can come PERF
 Bophaa can come.

In sentence (12), the subject has permission to perform a certain action, in this case to come. This permission can stem from either the speaker or a third person.

Thus, the verb *ʔaac* behaves not unlike English *can*. Just as is the case with *can*, *ʔaac* serves as auxiliary verb for both ability and permission. In actuality, then, both sentences (11b) and (12) are ambiguous between both readings. Sentence (11b) can be understood as permission to speak Khmer and (12) can be taken in the sense that the subject has the ability to come.

This draws to a close the section on deontic modality. We have seen that Khmer has modal verbs to express the same basic meanings as the corresponding English verbs. Turning to the epistemic continuum now, we see a different picture.

3.2 EPISTEMIC MODALITY

In Khmer, the epistemic continuum contains only one grammaticised element, namely the verb *kuo* 'should'. The degree of epistemicity expressed lies roughly between the weak and strong epistemic extremes. Unlike English, Khmer uses no verbs to express weak or strong epistemic modality. The first approximation of the Khmer continuum is shown in Chart 9a:

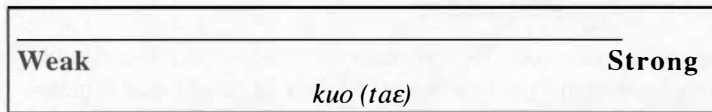


CHART 9a: KHMER EPISTEMIC MODALS: FIRST ATTEMPT

However, this cannot be the true representation of the Khmer epistemic modal system, since the representation in Chart 9a implies that the verb *kuo (taε)* covers the entire range of the epistemic spectrum, from weak to strong modality, which is emphatically not the case. In order to show that *kuo (taε)* has a limited range, I will use a dotted line for those parts that are not represented by grammaticalised forms. This is shown in Chart 9b:

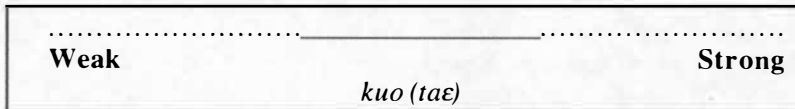


CHART 9b: KHMER EPISTEMIC MODALS: FINAL REPRESENTATION

The representation of Chart 9b shows that the verb *kuo (taε)* has a limited range, but it also takes into account the fact that the boundaries between the areas where *kuo(taε)* is appropriate and where it is not are fuzzy.

Neither the verb *trəw* 'must' nor *ʔaac* 'can, may' can be used in an epistemic sense. The verb *trəw* is used only deontically and *ʔaac* is used for deontic modality and for ability, but not for epistemic modality. For strong and weak epistemic modality, recourse must be sought in adverbs, and adverbs do not have a place on the modal continuum. As explained in §2 above, I have chosen to put only grammaticised meanings on the continuum.

We have already encountered examples of sentences with the verb *kuo* in (9) and (10) above. In those instances, *kuo* served as a verb expressing deontic modality. Sentence (13) is an example of a sentence with *kuo* in which the epistemic sense is the most natural interpretation, but it must be kept in mind that (13) is in reality ambiguous between epistemic and deontic modality, just as (9) and (10) above are. Although sentence (13) can be interpreted as expressing an obligation that the room be dark, the most logical interpretation is one of inference: based on known facts, the room is dark, although there is no verification of the statement. The speaker expresses a high probability that the event is true.

- (13) *Bantep kuo (taε) ngohngit.*
 room should be.dark
 The room should be dark.

Just as in its deontic interpretation, the verb *kuo taε* can be used instead of *kuo*. The main reason for the choice of *kuo taε* over *kuo* appears to be stylistic. For instance, my informant judged sentence (13) 'smoother' with *kuo taε* instead of *kuo*, but when a negation is present, the simple form *kuo* is preferred.

The Khmer modal system outlined so far is syntactically homogeneous. The Khmer modal verbs behave similarly with respect to each other. The last two modal forms discussed here are different from modal verbs syntactically. The forms that express weak and strong epistemic modality are not verbs, but adverbs.

3.2.1 THE ADVERB *prahæl*

Weak epistemic modality in Khmer is expressed by the adverb *prahæl* 'maybe'. The use is shown in (14) below. Sentence (14) denotes that the speaker thinks there is a possibility that the action expressed in the sentence is true. He is not committed to this possibility, however. It is equally possible that the action is not true. The sentence might be paraphrased as: 'It is possible that Bophaa killed the fish'.

- (14) *Bophaa prahæl (ciə) baan səmlap trəy.*
 Bophaa maybe be PERF kill fish
 Bophaa may have killed the fish.

A difference between modal verbs and adverbs is readily apparent from sentence (14): the presence of the copula *ciə* 'be'. It is true that the copula is optional, but the sentence without *ciə* is considered to be very awkward. The modal verbs discussed above can never be combined with the verb *ciə*.

Modal adverbs are also different from modal verbs in that they can be preposed more easily. Although sometimes modal verbs can be found in sentence-initial position, as seen in sentence (5), this is extremely rare and sentences like (5) are usually expressed as questions. Modal adverbs can easily be placed in first position without losing the declarative meaning:

- (15) *Prahæl (ciə) Bophaa baan səmlap trəy.*
 maybe be Bophaa PERF kill fish
 Bophaa may have killed the fish.

To form questions, question particles are used:

- (16) *Ta prahæl (ciə) Bophaa baan səmlap trəy rii?*
 Q maybe be Bophaa PERF kill fish Q
 Could Bophaa have killed the fish?

A possible answer to question (16) might be (17a), but not (17b):

- (17)a. *Prahæl.*
 maybe
 Maybe.
 b. **Bophaa prahæl.*
 Bophaa maybe

Sentence (17b) shows another difference between modal verbs and adverbs. It is not permissible for the response to a question to be just the subject plus the modal adverb. The

answer, however, can be the combination of the subject plus a modal verb. This is shown in (18):

- (18)a. *Bophaa ?aac səmlap trəy rii?*
 Bophaa can kill fish Q
 Can Bophaa kill the fish?
- b. *Bophaa ?aac.*
 Bophaa can
 (Yes,) Bophaa can.

3.2.2 THE ADVERB *pīt ciə*

The adverbial construction *pīt ciə* (the literal meaning of which is 'be right') is used to denote the epistemic notion of necessity. It can be used with animate as well as inanimate subjects:

- (19)a. *Bophaa pīt ciə baan səmlap trəy.*
 Bophaa right be PERF kill fish
 Bophaa must have killed the fish.
- b. *Bantep pīt ciə ngohngit.*
 room right be be.dark
 The room must be dark. (Surely, the room is dark.)

When a speaker uses *pīt ciə*, he is very sure of himself. Based on his knowledge of the situation, he is convinced that what he says is true. The only reason the speaker does not use a normal declarative sentence is in order to guard against the (very unlikely) possibility of the situation being different from what he believes it to be.

There are two main differences between the syntax of *pīt ciə* and the syntax of *prahæl* (*ciə*). The most obvious one is that the use of the copula in *pīt ciə* is obligatory, while it is optional in *prahæl* (*ciə*), even though it is considered to be awkward to omit the copula.

- (20) **Bantep pīt ngohngit.*
 room right be.dark

The second difference is the interpretation of preposed *pīt ciə*. Whenever *pīt ciə* is preposed, the sentence is interpreted as a question, just as if there were regular modal verbs present. This is shown in (21) below:

- (21) *Pīt ciə Bophaa baan səmlap trəy rii?*
 right be Bophaa PERF kill fish Q
 Is it true that Bophaa killed the fish?

A sentence with the adverb *prahæl* preposed is usually interpreted as a declarative sentence, as seen in sentence (15) above. From these observations, then, we can draw the conclusion that the Khmer epistemic system is heterogeneous in nature, and that even the distinction verb versus adverb is not enough of a distinction to account for the syntactic differences between the various modal elements. In this respect, Khmer is no different from English, where similar syntactic differences can be observed between modal verbs, adverbs, and other modal elements. English modal adverbs can be moved freely, just as Khmer modal adverbs can.

3.3 DIACHRONIC DEVELOPMENTS

Even though the present study deals mainly with the synchronic modal system of Khmer, a few words on the diachronic development seem in order. The fact that only one verb is present in the Khmer epistemic continuum is a confirmation of universal diachronic changes in modal systems. It has been observed that changes in the modal system go from an ability or deontic interpretation to an epistemic interpretation, not from epistemic to deontic or ability. Bybee et al. (1994, Chapter 6) present an impressive range of data from a well-balanced sample to support this hypothesis.

The modal system of English developed in this way. The verbs *can* and *may* were originally used to denote ability, but gradually acquired first deontic and then epistemic interpretations. The same can be said for *must*; it started out as a deontic verb, and then acquired epistemic interpretations as well. Finally, the verb *should* started out as the past tense of *shall*, but already in Old English it had acquired interpretations of its own (Visser 1969:1636ff). Even its epistemic interpretations seem to start in the Old English period (Traugott 1989:41), but the deontic interpretations seem to be most numerous.

In view of this, it seems likely that the Khmer modal verb *kuo* developed its epistemic interpretation after the deontic interpretation. Apart from the universal tendencies mentioned above, another argument for this development is the fact that there are no other verbs to denote weak and strong epistemic modality. It seems most likely that *kuo* is the first verb to cross the line between deontic and epistemic, because there is no real epistemic, adverbial counterpart for *kuo* in the language. The verb *ʔaac* has *prahæɭ* as counterpart, while *trəw* has *pʰi ciə*. The presence of modal adverbs to fill the epistemic gap is possibly the reason that both *ʔaac* and *trəw* have not developed epistemic interpretations.

The proposed analysis is then a broadening in meaning for *kuo* from deontic to epistemic, and a resistance (for the present) of *ʔaac* and *trəw* to do the same. Nevertheless, this analysis is speculative, and research in the history of Khmer is needed to either substantiate or refute the hypothesis.⁷

4. MODALITY AND NEGATION

We now turn our attention to the interaction between modality and negation. In principle, there is no reason why we should not establish a similar continuum for the negative elements, just as I have done for the modal verbs discussed in §3 above.⁸ In this paper, however, I will take a different approach. I feel that a continuum model is not really suited to show the

⁷ In order to convincingly explain the situation it is also important to examine the modal systems of surrounding languages. In Thai and Lao (languages not related or not closely related to Khmer), for instance, the verb denoting weak modality is *aat*, clearly a cognate of Khmer *ʔaac*. In Lao, *aat* has both epistemic and deontic readings (Bybee et al. (1994:249), citing Steele 1975), but in Thai, *aat* has only an epistemic interpretation. Matisoff (1991:393) gives an example of Thai *aat* in an ability reading. The Lao data can be explained by means of the extension of meaning from deontic to epistemic modality but the Thai data are puzzling. If *aat* in Thai has only an epistemic and an ability reading, we might be looking at a counterexample to the universal given above.

⁸ For an approach along these lines see de Haan (1989), in which the modal systems of Russian and Chinese were compared in this way with the category of negation.

possible interactions of modality and negation. Instead, I will use a framework I developed elsewhere (de Haan 1994) to describe the typological differences between modal verbs and negation in various languages. The framework is briefly sketched here. For more details, see de Haan (1994).

4.1 BASIC FRAMEWORK

Modality and negation each interact with the sentence in which they appear. If a modal element is added to a sentence (p), then the resulting sentence MOD(p) reflects a certain attitude on the part of the speaker with respect to (p). Similarly, to place a negation in (p) results in changing the truth value of the sentence. If (p) is true, then NEG(p) is false, and vice versa. Of course, there are many kinds of negation (see Horn (1989) for discussion). The kinds I am concerned with here are sentence and constituent negation, following Klima (1964). I take negation to be a syntactic category here.

If both modality and negation are present, they not only interact with p, but also with each other. This is shown in (22) with examples from English. Following each sentence is the representation I will use in this paper. It should be kept in mind that I am not dealing with logical relations here, only with linguistic representations. These two domains should be kept separate, hence the use of MOD and NEG instead of the more usual logical symbols.

- (22)a. John goes to school. (p)
 b. John must go to school. (MOD (p))
 c. John doesn't go to school. (NEG (p))
 d. John mustn't go to school. (MOD (NEG (p)))
 e. John needn't go to school. (NEG (MOD (p)))

The crucial sentences are (22d) and (22e). The difference between the two sentences is one of **scope**. In sentence (22d) the negation is in the scope of the modal, while the modal is in the scope of the negation in (22e). In sentence (22d), the negation has **narrow scope**; in (22e), it has **wide scope**.

In this section I will also make a difference between **basic notions** and **instances** of that notion. An example of a basic notion is strong deontic modality, represented by MUST. A basic notion is graphically represented by means of small caps. In English, this basic notion is instanced by means of the verb *must*. Instances are represented in this section in italics. At first, it may seem confusing to have these basic notions labelled by the names of English verbs, but this is done to facilitate recognition of the notion. Instead of using these labels, we could have used terms such as 'strong deontic modality' (for MUST) but it is more convenient to use the label. Moreover, for some notions, there is no convenient term available, such as for the notion SHOULD. The reader needs to distinguish between MUST, a cross-linguistic category, from *must*, its reflex in English. When dealing with other languages, the apparent confusion diminishes. In Khmer, for instance, the notion MUST is instanced by means of the verb *trəw*.

In de Haan (1994), I investigated how the difference in scope is represented in various languages. There are basically two types of strategies that play a role. The first one, called the Modal Suppletion Strategy, or MSS for short, is characterised by the use of a different modal verb is used to make the difference in scope. The MSS is formalised in (23) below:

- (23) Modal Suppletion Strategy
- a. Neg V_{mod1} V_{main} (MOD (NEG (p)))
 - b. Neg V_{mod2} V_{main} (NEG (MOD (p)))

It must be kept in mind that the linear order of the elements is irrelevant, although a correlation between relative order of NEG and MOD and basic word order exists, as shown in de Haan (1994). In some languages, negation may precede the modal verb (Finnish is such a language), while in others it may follow it (as seen in (22), with English as an example). What is relevant is whether a different verb is used for the difference in scope. The negative elements in (23a) and (23b) behave the same, as far as the syntax is concerned.

The second strategy is called the Negation Placement Strategy, or NPS. It is characterised by the fact that the place of the negation changes to make changes in scope. The NPS is formalised in (24):

- (24) Negation Placement Strategy
- a. Neg V_{mod} V_{main} (NEG (MOD (p)))
 - b. V_{mod} Neg V_{main} (MOD (NEG (p)))

Again, linear order of the various elements is not really relevant. All that matters is that there are two different places in the sentence for the negation to surface. The modal verb is the same in both (24a) and (24b). Crucial is that the negation has different syntactic properties in these instances. These properties can be best shown by applying the Klima tests, as originally developed for English in Klima (1964). The Klima tests consist of a series of tags that can be added to matrix sentences containing a negative element but cannot be added to positive sentences. These tags are:

- (25)a. a tag with *neither*.
 b. a tag with *not...either*.
 c. a tag with *not even*.
 d. a question tag of opposite polarity.

The negation in (24a) will be called SENTENCE NEGATION; in (24b), it is called CONSTITUENT NEGATION. The difference between sentence and constituent negation is that tag clauses can be added to clauses with sentence negation, but not to clauses with constituent negation. For an example from English see (35) below. Examples of other languages that make use of NPSs are Italian (and Romance languages in general), Russian, and Modern Greek.

Another process that plays an important role with NPSs is that of Negative Transportation (NT). NT can be characterised by the fact that the structural representation of wide scope (24a) can be interpreted semantically as narrow scope. Thus, NT is a process that creates ambiguities. The representation of (24a) has two interpretations: the negation can have both wide and narrow scope. The representation of (24b) is still unambiguously an instance of narrow scope. This process also plays an important role in Khmer.

4.2 THE INTERACTION OF NEGATION AND MODAL VERBS IN KHMER

Khmer is interesting in that it seems to be a language with both MSSs and NPSs. The start of the discussion will be the modal verb *trəw* 'must'.

In §3.1, I discussed the behaviour of negation with the verb *trəw*. We saw that the negation particle *mɪn* is placed before the modal verb to express narrow scope of the negation (sentence (8a), repeated below as (26a)). In addition, there is a verb (*cam*) *bac*, which in combination with the negation particle is used for wide scope of the negation (sentence 8b). In this instance, Khmer makes use of the Modal Suppletion Strategy.

There is reason to believe that this is not the only possible way, however. It was already remarked in the discussion of (8) that the negative particle can also be placed after the modal verb *trəw*. This is shown in (26b) below. Unlike sentence (26a), (26b) is unambiguous in its interpretation. The representation is analogous to (24b), and the negation can be interpreted only as having narrow scope. The process of NT is responsible for the scope ambiguities of sentence (26a). Sentence (8b) is repeated here as (26c).

- (26)a. *Kñom mɪn trəw səmlap trəy.*
 I NEG must kill fish
 I mustn't/needn't kill the fish.
- b. *Kñom trəw mɪn səmlap trəy.*
 I must NEG kill fish
 I must not kill the fish.
- c. *Kñom mɪn cam.bac səmlap trəy.*
 I NEG need kill fish
 I needn't kill the fish.

Khmer can then be analysed as having both MSSs and NPSs. The contrast of the sentences shown in (26a and 26c) are instances of the MSS, while the contrast shown in (26a and 26b) points to the NPS. This is a curious situation. Although it is not uncommon for a language to have both MSSs and NPSs (as is evidenced by English; see §5), it is uncommon to see MSSs and NPSs duplicate each other's functions, as Khmer does. It is quite uneconomical to have different expressions for the same semantic load, especially in the strong deontic part of the continuum. The Khmer data are summarised in (27).

- (27) Strong deontic Modality and Negation in Khmer
- | | | | |
|-----------------|----------------------|------------|---------|
| (MOD (NEG (p))) | <i>trəw mɪn</i> | 'must not' | (NPS) |
| | <i>mɪn trəw</i> | | (by NT) |
| (NEG (MOD (p))) | <i>mɪn trəw</i> | 'need not' | |
| | <i>mɪn (cam) bac</i> | | (MSS) |

The other modal verbs, *ʔaac* 'may, can' and *kuo* 'should', can also occur with the negation particle *mɪn* placed either before or after the modal. In the case of *ʔaac*, the placement of *mɪn* after it (sentence (28b)) results in the interpretation of (24b); the negation has narrow scope, that is a permission **not** to perform a certain action:

- (28)a. *Bophaa mɪn ʔaac səmlap trəy.*
 Bophaa NEG can kill fish
 Bophaa cannot kill the fish. (NEG (MOD (p)))
 (MOD (NEG (p)))
- b. *Bophaa ʔaac mɪn səmlap trəy.*
 Bophaa can NEG kill fish
 Bophaa can not kill the fish. (MOD (NEG (p)))

Sentence (28a) reflects the normal state: the negation is placed before the modal verb and the sentence is ambiguous in the interpretation of the scope of the negation. Both interpretations (wide and narrow scope) are accepted. The wide scope interpretation is a reflection of the linear order of negation and modal verb (see (24a)), while Negative Transportation causes the narrow scope interpretation.

The negative possibilities of the verb *kuo* 'should' are interesting, since *kuo* is placed roughly in the middle of the modal continuum (see Chart 7 and Chart 9b for the deontic and epistemic continuum, respectively). In English, when a negation is combined with *should*, the negation always has narrow scope. This is no doubt due to the fact that *should* derives from the stronger modal *shall*. A negation combined with *shall* also has narrow scope. Recall also the discussion in §2 of the Dutch translation of the notion of *should*: the Subjunctive of the strong modal *moeten*. The verb *moeten* also causes a negation to have narrow scope.

Since notions such as SHOULD have only one scope interpretation when combined with a negation, we would expect Khmer *kuo* to have only one possible place for the negation, but this turns out not to be the case. The verb *kuo* allows a negative element to be inserted into both possible slots:

- (29)a. *Kñom mɨn kuo səmlap trəy.*
 I NEG should kill fish
 I shouldn't kill the fish. (MOD (NEG (p)))
- b. *Kñom kuo mɨn səmlap trəy.*
 I should NEG kill fish
 I shouldn't kill the fish. (MOD (NEG (p)))
- (30)a. *Bantep mɨn kuo ngohngit.*
 room NEG should be.dark
 The room should not be dark. (MOD (NEG (p)))
- b. *Bantep kuo mɨn ngohngit.*
 room should NEG be.dark
 The room should not be dark. (MOD (NEG (p)))

Even though the negation can be placed either before the modal (sentences (29a), (30a)) or after it ((29b), (30b)), there is no difference in interpretation. Both sentences are interpreted as the modal verb having scope over the negation. There is no interpretation available in which the negation has scope over the modal verb. It makes no difference whether the verb *kuo* is interpreted as deontic (29) or epistemic (30). In both cases the only interpretation of the sentence is that in which the modal has scope over the negation.

There seems to be no difference, then, in the interpretation of the notion SHOULD between languages that make use predominately of MSSs and languages that make use mostly of NPSs. In the first case, there is no verb corresponding to *should* to express the notion of SHOULD plus a negation that has wide scope over it. MSSs seem to be limited to modal verbs that express weak and strong Modality, and not the in-between points on the continuum. In the case of NPSs, there is simply no difference in interpretation between the syntactic structures of (24a) and (24b). Both are interpreted semantically as narrow scope.

Both on my continuum model and on the more usual scalar models, the notion of SHOULD is situated between weak and strong Modality. Weak and strong modality can be logically expressed in terms of each other. This is shown in (31), by the use of a representative of

both strong and weak modality. The symbol 'nec' refers to 'necessity', that is to strong modality, and 'poss' refers to 'possibility', that is to weak modality. The formulas work equally well with other representatives of strong and weak modality, for instance with obligation and permission, respectively. These logical equivalents are well known. See, for example, Lyons (1977:165) for a discussion.

- (31)a. nec p = not poss not p
 b. poss p = not nec not p

Also, negation of necessity and possibility can be expressed in terms of each other:

- (32)a. not nec p = poss not p
 b. not poss p = nec not p

In logical terms, necessity and possibility can be defined in terms of each other. This is not the case with *should*. *Should* cannot be expressed in terms of another modal operator. Indeed, in modal logic, the notion of SHOULD does not have a place at all. It comes as no surprise, then, that negation combined with such an element has only one interpretation, no matter whether the negation is placed before or after the modal.

5. CONCLUSIONS

In this paper, I contrasted the modal systems of Khmer and English. Even though both languages make use of modal verbs and a separate morpheme for negation, the way modal notions and the interaction of modality and negation are expressed differ radically.

In English, deontic and epistemic notions are encoded by means of the same verbs. For this reason, the epistemic modal continuum in English is very similar to that of the deontic continuum. In Khmer on the other hand, the basic weak and strong epistemic notions are expressed by means of modal adverbs, rather than modal verbs. Consequently, the epistemic continuum (see Chart 9b) is emptier than the deontic one (see Chart 7).

It is true that the situation in English seems to be the preferred situation from a typological point of view. The notions of deontic and epistemic modality seem to be much more often expressed by the same elements than by different elements. For instance in Tamil (Asher 1982:167–172), strong modal notions, both deontic and epistemic, are expressed by the suffix *-ñum*, while weak modal notions are expressed by means of *-laam*.

However, the Khmer situation is not unique. Palmer (1986:123) cites the example of literary Arabic, which has four distinct forms:

- | | | | |
|------|-----------|------|----------------|
| (33) | Epistemic | MAY | <i>rubbama</i> |
| | | MUST | <i>labudda</i> |
| | Deontic | MAY | <i>yumkin</i> |
| | | MUST | <i>yagib</i> |

As noted earlier, the situation in Khmer can be explained only against the background of other Southeast Asian languages, in particular Thai and Lao, where the distribution of the modal notions is different from Khmer, but where the verbs used are cognates of the Khmer modal verbs.

As far as negation is concerned, we saw that Khmer and English have two different basic patterns; Khmer is basically an NPS language, while English is basically an MSS language.

As we have also seen, Khmer does possess one MSS strategy, namely the verb *cam bac*. In contrast, English has an NPS strategy in the verb *may, might, can* and *could* when combined with a negation. This can be seen in (34) below. The occurrence of *may not* in (34a) is an example of an MSS strategy, since the modal is in the scope of the negation. It can be paraphrased by *You are not allowed to go to the movies tonight*. In (34b), however, it is an NPS strategy. The negation is in the scope of the modal, as witnessed by the paraphrase *It is possible that it won't rain tomorrow*.

- (34)a. *You may not go to the movies tonight.*
 b. *It may rain or it may not rain tomorrow.*

The difference between the two strategies can be seen most clearly, when we combine them with a Klima-tag (see (25) above). The fact that sentence (34a) can be combined with a Klima tag, as shown in (35a) below, while (34b) cannot, as seen in (35b), shows that the negation in (34b) is an example of a constituent negation, while the negation in (34a) is sentential. This is the distinctive characteristic between MSS and NPS strategies, as argued briefly above and in de Haan (1994) in more detail. Note that sentence (35b) below is only ungrammatical in the narrow scope reading of the negation. It is, of course, perfectly grammatical in the wide scope, or prohibition, reading.

- (35)a. *You may not go tonight, and neither may Mary.*
 b. **It may not rain tomorrow, and neither may it snow.*

The pattern found here seems to be valid cross-linguistically. If a language makes for the most part use of MSS strategies, it can have NPS strategies in the weak epistemic and deontic part of the continuum. This is true in other Germanic languages (though not in Finnish, another MSS language). If, on the other hand, a language makes predominant use of NPSs, it may have an MSS strategy in the strong deontic part of the continuum. This is not only exemplified by Khmer, but also by Spanish (*necesitar*) and Modern Greek (*xreiazomai*).

REFERENCES

- Asher, R.E., 1982, *Tamil*. Lingua Descriptive Studies 7. Amsterdam: North-Holland Publishing Company.
- Boyd, J. and J. P. Thorne, 1969, The semantics of modal verbs. *Journal of Linguistics* 5:57–74.
- Bybee, Joan, Revere Perkins and William Pagliuca, 1994, *The evolution of grammar: tense, aspect and modality in the languages of the world*. Chicago: University of Chicago Press.
- Coates, Jennifer, 1983, *The semantics of the modal auxiliaries*. London: Croom Helm.
- De Haan, Ferdinand, 1989, Modaliteit in twee werelden: een onderzoek naar modale hulpwerkwoorden in het Russisch en Chinees. Unpublished MA thesis, Groningen University.
- 1992, Modality in Dutch. MS, University of Southern California.
- 1994, The Interaction of modality and negation: a typological study. PhD thesis, University of Southern California.
- Ehrman, Madeline, 1966, *The meaning of the modals in present-day American English*. The Hague: Mouton.
- Hengeveld, Kees, 1987, Clause structure and modality in Functional Grammar. In Johan van der Auwera and Louis Goossens, eds *Ins and outs of the predication*, 53–66. Dordrecht: Foris.
- Horn, Laurence R., 1972, On the semantic properties of logical operators in English. PhD thesis, University of California at Los Angeles. Distributed by the Indiana University Linguistic Club, 1976.
- 1989, *A natural history of negation*. Chicago: University of Chicago Press.
- Jacob, Judith M., 1968, *Introduction to Cambodian*. Oxford: Oxford University Press.
- Klima, Edward S., 1964, Negation in English. In J.A. Fodor and J.J. Katz, eds *The structure of language: readings in the philosophy of language*, 246–323. Englewood Cliffs, NJ: Prentice-Hall.
- Lyons, John, 1977, *Semantics*. Cambridge: Cambridge University Press.

- Matisoff, James A.**, 1991, Areal and universal dimensions of grammatization in Lahu. In Bernd Heine and Elizabeth Closs Traugott, eds *Approaches to grammaticalization*, 383–453. Amsterdam: John Benjamins.
- Palmer, Frank R.**, 1986, *Mood and modality*. Cambridge: Cambridge University Press.
- 1990, *Modality and the English modals*. London: Longmans.
- Siewierska, Anna**, 1991, *Functional grammar*. London: Routledge.
- Steele, Susan**, 1975, Is it possible? *Stanford working papers in language universals* 18:35–58.
- Traugott, Elizabeth Closs**, 1989, On the rise of epistemic meaning: an example of subjectification in semantic change. *Language* 65:31–55.
- Twaddell, W.F.**, 1960, *The English verb auxiliaries*. Providence, RI: Brown University Press. (2nd, revised, edn 1963.)
- Visser, F.Th.**, 1969, *An historical syntax of the English language, Part III, first half: syntactical units with two verbs*. Leiden: E.J. Brill.

SOCIO-SEMANTIC ASPECTS OF HUMAN MEASURE WORDS IN CANTONESE

JUDY HO

1. INTRODUCTION

Measure words are an interesting phenomenon which reveals a great deal about the cognitive organisation of the world of the people who use them. (Craig 1986; Lakoff 1986; Dixon 1986). Some languages have highly developed classifier or measure-word systems, which have been studied by a number of scholars. (See, for example, Adams 1986, 1989 in Austroasiatic, Becker 1975 in Burmese, Juntanamalaga 1988 in Thai and Denny 1979 in Japanese.) Some of those who use measure words make up a quarter of the world's population. Here I am referring to the speakers of Chinese. However, although much enthusiasm has been evinced in the study of Chinese, most of the research on the properties and uses of measures has been carried out on Mandarin only. (See, for example, Chao 1968, DeFrancis 1984, Erbaugh 1986, Ramsey 1987 and Norman 1988). What I propose to do in this paper is to examine some measure words for humans in the dialect of Hong Kong Cantonese with regard to their semantic properties and pragmatic, social and functional roles, as well as their relationship to the nature of categorisation in a given culture.¹

I have chosen Hong Kong Cantonese for this study for two reasons: First, as stated above, comparatively little research has been done on measure words in Cantonese, which is the native language of about five per cent of the Chinese-speaking population. It is the principal Chinese dialect used in Hong Kong, and many overseas Chinese in Australia, the U.S.A., Canada, the U.K., Singapore, Malaysia and Vietnam are also native speakers of Cantonese.²

The second and a more important reason is that, given its political, social and economic structure, Hong Kong is a predominantly Chinese society open to Western influences in thinking, technology and other areas. This sociocultural uniqueness of Hong Kong offers interesting opportunities to study semantic shifts and motivation for the extension of categories.

¹ The author is grateful to Prof. Ron Scollon for his valuable comments and assistance in the preparation of this paper.

² The information about Cantonese speakers is based on Ramsey (1987) pp.87, 98–99, and on NAATI 1987 Candidates' Manual p.30.

2. OVERVIEW

Measure words are an important syntactic as well as semantic feature in Cantonese. A numeral cannot directly modify a noun except in idioms, proverbs or literary writing. Therefore example (1a) is ungrammatical while (1b), with the measure *jék*, is grammatical.³

- (1) a. ㄣ 〇 ㄣ
 ? *Yau* *yat* *ma*.
 have one horse
 There is a horse.
- b. ㄣ 〇 ㄣ ㄣ
 Yau *yat* *jék* *ma*.

However, it is perfectly acceptable to use an idiom which does not have the classifier, as in (2).

- (2) 〇 〇 ㄣ ㄣ
 saam *sam* *leung* *yi*
 three heart two will
 a very undecided person

Another thing which is worth noting is that when Cantonese speakers refer to a noun at the generic level, they do not use any measure words. For example:

- (3) ㄣ ㄣ ㄣ 〇 ㄣ ㄣ
 Yan *hai* *ho* *chung* *ming* *ge*.
 person is very clever bright MOD
 Human beings are very intelligent.
- (4) ㄣ ㄣ ㄣ ㄣ ㄣ ㄣ ㄣ
 Gau *hai* *yau* *yung* *ge* *dung* *mat*.
 dog be have use MOD moving thing
 Dogs are useful animals.

In (3) and (4), we find that reference to human beings or animals as a species does not require measure words. However, if the referents are individual entities of a species as in (5) and (6), measure words (MEA) are required.

- (5) 〇 ㄣ 〇 ㄣ ㄣ ㄣ
 yat *goh* *chung* *ming* *ge* *yan*
 one MEA clever bright MOD person
 a clever person
- (6) 〇 ㄣ ㄣ ㄣ ㄣ ㄣ
 yat *jék* *yau* *yung* *ge* *gau*
 one MEA have use MOD dog
 a useful dog

³ The phonetic symbols and tone marks used in this paper are adapted from Lam and Po's (1988) *Functional Cantonese Book One*, pp.vii–xi for their simplicity, clarity and visual appeal. Although there are 9 tones in Cantonese, only 6 basic tones on a scale of 5 pitch levels are represented here. The remaining 3 tones are the so-called 'entering tones' which are not unique tones because their musical values are identical to the upper level, upper going and lower going tones respectively (S.L.Wong 1995:106).

3. CANTONESE ANIMATE MEASURE WORDS

Measures not only carry syntactic significance; they also possess interesting semantic features, the discussion of which will be based on the individual measures illustrated in the following figure:

human beings	animals	birds	fishes	insects
\bar{goh}	\bar{jek}	\bar{jek}	$\bar{tiù}$	\bar{jek}
$\bar{wái}$	$\bar{tiù}^*$		$\bar{dâu}^*$	$\bar{tiù}$
$\bar{tiù}$	$\bar{pāt}^*$			
\bar{dau}				
\bar{jek}				

* = specific use

FIGURE 1: MEASURE WORDS FOR ANIMATE BEINGS IN CANTONESE

The figure shows five categories of living creatures: human beings, animals, birds, fishes and insects. The data have been selected to show the complexity of the system of Cantonese measure words. Some measure words are unique to a particular category, while others are applicable to more than one category; for example, $\bar{wái}$ is a measure word for human beings only, whereas \bar{jek} can be used to modify animals, birds, insects and even human beings in particular contexts. I propose to begin with the fairly simple semantic elements pertaining to measures used for non-human living creatures and then proceed to the more sophisticated use of measures for human beings.

3.1 MEASURES FOR NON-HUMAN BEINGS

One of the basic principles of categorisation in Cantonese measure words (as in many other noun classifiers) is the human cognitive representation of the shape, size, and attributes of the creature or object to be modified.

The measure \bar{jek} is used to modify any creatures that have legs, so the following expressions are used:

- (7) \bar{yat} \bar{jek} \bar{gau}
 one MEA dog
 a dog
- (8) \bar{yat} \bar{jek} \bar{ying}
 one MEA eagle
 an eagle

Insects that are seen more often to crawl on the ground with legs are modified by \bar{jek} too, as in (9) and (10).

- (9) \bar{yat} \bar{jek} $\bar{ji-jue}$
 one MEA spider
 a spider

- (10) $\overset{\circ}{y}at$ $\bar{j}ek$ $\bar{g}aat-\check{j}aat$
 one MEA cockroach
 a cockroach

Even though cockroaches have wings and are able to fly, their crawling with legs is seen by the human eye to be a more prominent attribute. Therefore $\bar{j}ek$ is used.

The measure $\grave{t}iu$ is used to modify creatures that are long in shape and have no legs, as in (11)–(13).

- (11) $\overset{\circ}{y}at$ $\grave{t}iu$ $\acute{y}ue$
 one MEA fish
 a fish

- (12) $\overset{\circ}{y}at$ $\grave{t}iu$ $\grave{c}hung$
 one MEA worm
 a worm

- (13) $\overset{\circ}{y}at$ $\grave{t}iu$ $\grave{s}e$
 one MEA snake
 a snake

As in the case of cockroaches, crawling is a prominent attribute of worms and snakes, but the measure word $\bar{j}ek$ is not used because the latter do not crawl with legs.

Another interesting example which shows the importance of shape in determining measure words for non-human creatures is:

- (14) $\overset{\circ}{y}at$ $\grave{t}iu$ $\grave{k}ing$ $\grave{y}ue$
 one MEA whale fish
 a whale

Although whales are biologically classified as mammals, they are modified by $\grave{t}iu$ (which is the measure word for fish), and not $\bar{j}ek$ (which is the measure word for mammals) because the shape of whales is seen to be fish-like.

Now let us look at those measure words marked with specific use.

In Cantonese, 'dog' would be the only animal modified by $\grave{t}iu$. Even so the measure is rarely used in the spoken form of Cantonese. It is sometimes used in writing, which shows the influence from Mandarin as in (15).

- (15) $\bar{y}i$ $\acute{t}iao$ $\check{g}ou$
 one MEA dog
 a dog

Example (15) is a common expression in the northern dialects of China. It should be pointed out that Cantonese speakers write in Mandarin although they pronounce the written words in a totally different way from Mandarin speakers.

$\overset{\circ}{P}at$ is another measure word which has very restrictive use. It modifies horses, cloth and waterfall only. The semantic link between these three objects is probably the rolling motion of racing horses, heavy rolls of cloth being unfolded and running waterfalls.

We can see the use of *pāt*^o to modify a horse in (16).

- (16) *yāt*^o *pāt*^o *mā*[┘]
 one MEA horse
 a horse

However, when this measure is used in the spoken form in Hong Kong, it is normally used in a specific context, that of horseracing commentaries. When used in writing, it is a slightly more formal word and more of a literary style than *yāt jek mā*^o.

The measure *dāu*^o has a use similar to *tiù*^o and modifies fishes only. It is less often used than *tiù*^o and is a dialectal variation.

3.2 MEASURE WORDS FOR HUMANS

As we turn to the measures used for human beings, we notice in them an intertwining of semantic, pragmatic and social roles. The most neutral measure is *gōh*^o, which is used to refer to a human person, regardless of the person's sex, age,

occupation or any other attributes. Examples (17)–(19) show that *gōh*^o is not sex specific, and (18)–(21) show that it does not make age distinctions.

- (17) *yāt*^o *gōh*^o *yān*[┘]
 one MEA person
 a person
- (18) *yāt*^o *gōh*^o *naam*[┘] *yān*[┘]
 one MEA male person
 a man
- (19) *yāt*^o *gōh*^o *nui*[┘] *yān*[┘]
 one MEA female person
 a woman
- (20) *yāt*^o *gōh*^o *naam*[┘] *jai*[┘]
 one MEA male small
 a boy
- (21) *yāt*^o *gōh*^o *nui*[┘] *jai*[┘]
 one MEA female small
 a girl

The word *jai*[┘] literally means 'son' but in Cantonese it means 'small' when it functions as a diminutive suffix and is similar to '-let' in English words such as 'streamlet' and 'booklet'.

Examples (22)–(24) show that *gōh*^o is used for people engaged in different jobs or with different educational backgrounds and social statuses.

- (22) *yāt*^o *gōh*^o *gung*^o *yān*[┘]
 one MEA work person
 a worker

- (23) $\overset{\circ}{y}at$ $\bar{g}oh$ $\bar{h}aa\bar{u}$ $\acute{j}eung$
 one MEA school senior
 a school principal

- (24) $\overset{\circ}{y}at$ $\bar{g}oh$ $\acute{j}ung$ $\bar{l}ei$
 one MEA general manage
 a prime minister

Examples (25)–(27) demonstrate that $\bar{g}oh$ can be used to modify people with different outward appearances (beauty, tallness, etc.), inner qualities (intelligence, egoism, materialism, etc.) or other attributes.

- (25) $\overset{\circ}{y}at$ $\bar{g}oh$ $\bar{l}eng$ $\acute{n}ui$
 one MEA beautiful girl
 a beautiful girl

- (26) $\overset{\circ}{y}at$ $\bar{g}oh$ $\overset{\circ}{g}o$ $\acute{l}o$
 one MEA tall chap
 a tall chap

- (27) $\overset{\circ}{y}at$ $\bar{g}oh$ $\bar{s}ai$ $\bar{g}aci$ $\acute{j}ai$
 one MEA life boundary son
 a son of the world, a man or boy full of worldly wisdom

We have seen how $\bar{g}oh$ has the most unrestrictive register and is the most widely used measure word for human individuals.

By contrast, $\acute{w}ai$, which is also used for individual persons, has a far more restricted register. It usually indicates the speaker's courtesy and respect for the referent. The noun that it modifies is usually one that deserves or demands respect, as in (28)–(31).

- (28) $\overset{\circ}{y}at$ $\acute{w}ai$ $\bar{n}a\bar{m}$ $\bar{s}i$
 one MEA male high.status
 a gentleman

- (29) $\overset{\circ}{y}at$ $\acute{w}ai$ $\acute{n}ui$ $\bar{s}i$
 one MEA female high.status
 a lady

- (30) $\overset{\circ}{y}at$ $\acute{w}ai$ $\bar{j}ing$ $\acute{f}oo$ $\overset{\circ}{g}woon$ $\acute{y}uen$
 one MEA political office official member
 a government official

- (31) $\overset{\circ}{y}at$ $\acute{w}ai$ $\bar{l}o$ $\acute{y}an$ $\overset{\circ}{g}a$
 one MEA old person family
 an elderly person

In (28) and (29) $\acute{w}ai$ is used because it is an appropriate word to match 'gentleman' and 'lady', which are polite forms of address and reference, whereas in (30) a government

official demands respect by virtue of his official position and in (31) an elderly person deserves respect by virtue of his age.

However, in colloquial Cantonese, it is more common to use *gōh* in place of *wái* for the above examples. When *wái* is used, it carries with it a certain degree of formality and sheds light on the tone of the speaker and his relationship to the referent and to the hearer. A restaurant waiter will probably say to a customer:

- (32) *Gei* *doh* *wai*?
 how many MEA
 A table for how many people?

By using the measure word *wái*, the waiter is conscious of and intends to make others conscious of the refined and respectful language that he is using. The use of *gōh* instead of *wái* in this situation would sound a bit rude. But the same waiter would not sound rude if he says the following to his workmate:

- (33) *Hai* *goh* *dō* *yau* *saam* *goh* *haak*.
 in that place have three MEA customer
 There are three customers over there.

It should also be noted that *gōh* and *wái* are sometimes interchangeable but not always. It would be acceptable to use *gōh* for a lady, as in (34), the reason being that *gōh* is a neutral and general word.

- (34) *yāt* *goh* *nui* *si*
 one MEA female high.status
 a lady

However, it would be odd to say

- (35) ? *yāt* *wai* *nui* *yan*
 one MEA female person
 a woman

The measure *wái* is a more appropriate measure word for a lady and so is *gōh* for a woman. Moreover, *wái* definitely cannot be used to modify a beggar, thief or rascal, the outcasts of society, given the chasm between the scornful attitude that society normally shows towards these classes of people and the respect and courtesy that *wái* entails. Thus the distinctive uses of *gōh* and *wái* indicate the important pragmatic and social roles which Cantonese measure words play. This is similar to Craig's (1986:270) observation that Jacaltec noun classifiers vary "in accordance to the speaker's relationship to and attitude toward the person being talked about".

3.3 SOME SHIFTING MEASURE WORDS

In sharp contrast to the use of *wái*, *tiù* is never used in public speech because the language is considered vulgar. The measure *tiù* is a common measure word for both objects as in (36), and living creatures that are long in shape, as in examples (11)–(13) above.

- (36) ^oyāt tiù pèi daí
 one MEA leather belt
 a leather belt

The semantic shift in *tiù* needs to be studied in the particular social and cultural contexts of Hong Kong. The 1970s was a time of rapid economic and technological development in the colony. As a result of such development, television became an increasingly important part of ordinary people's lives. The influence of this mass medium was manifested in various areas of society, notably in the language use of the speech community. This background provided much motivation for the extension or crossover of categories. Since there was a trend towards using authentic language in dramatisation, vulgar language was used in stories depicting triad⁴ or criminal activities. Real or imagined language that a triad member was expected to use was imitated on television and radio. It was in this situation that phrases such as the following were first used:

- (37) ^uYau tiù ^uyau.
 have MEA bloke
 There is a bloke.

- (38) ^oyāt tiù ^uyau
 one MEA bloke
 a bloke

Example (38) can also be used as an adverbial phrase, meaning 'alone', or 'all by myself/himself', and so on.

The original purpose of using *tiù* was to express the speaker's strong emotion of anger, hatred or contempt for the referent by dehumanising him, by using a measure word that normally modifies non-human creatures and objects.

Over a period of ten years, the sense of *tiù* has gradually broadened to carry another connotation, that of people living illicitly as husband and wife, or their promiscuous behaviour. Because of this change in sense, two measure words are given another important role: they determine the meaning of the sentence, as shown in the following examples:

- (39) ^uKui ^uyau tiù ^unúi.
 3SG have MEA daughter
 He has a de facto wife.

- (40) ^uKui ^uyau ^ogoh ^unúi.
 3SG have MEA daughter
 He/She has a daughter.

- (41) ^uKui ^uyau tiù ^ujái.
 3SG have MEA son
 She has a de facto husband.

⁴ Triad societies are illegal organisations of gangsters. Their members are involved in a wide range of criminal activities but sadly many of the stories depicting notorious triad figures are given a positive and heroic interpretation in TV series and films.

- (42) $\overset{\cup}{K}ui$ $\overset{\cup}{y}au$ \overline{goh} $\overset{\prime}{j}ai$.
 3SG have MEA son
 He/She has a son.

So far I have analysed three clearly divergent senses of the word *tiù*: the first sense is based on human perception of the shape of the referent, the second is largely dominated by the speaker's attitude, and the third demonstrates social relationships. In fact, *tiù* also provides information about the speaker's social background.

The 1980s saw an increasing acceptance of the second sense of *tiù* (i.e. the sense of examples (37) and (38) into everyday use by people other than real triad members and criminals or dramatis personae). Now more and more members of the working class and people at the lower end of the social ladder have incorporated this word into their speech. Even a large number of male speakers who have a higher social status seem to find it particularly expressive. They tend to use it occasionally in private conversations with close friends in order to express their intense feelings, or indicate their masculinity in using strong language or just as a gesture of solidarity with other male speakers. Female speakers seldom use this word because it is overtly stigmatised language, although it may enjoy covert prestige among the groups of people mentioned above. The measure used in the third sense, that of de facto relations and promiscuous behaviour, was and still is considered a much coarser term. So when the word *tiù* is used to describe human persons, it generally reflects the low social class and personal qualities of the speaker; however, there is one exception, which will be discussed in the following paragraph.

The measure *tiù* modifies only a few nouns that refer to humans. The nouns *yáu*, 'bloke', *núi*, 'daughter, female', and *jái* 'son' as given in (37), (39) and (41) are the most common nouns that it qualifies. It is also used for the noun *hōn*, which means 'man' with its associations of manliness, physical strength, truthfulness, bravery, and so on. When *tiù* is used to modify *hōn*, the vulgarity pertaining to the measure word is immediately eliminated. For example, phrases such as in (43) and (44) are idiomatic expressions which were first used in literary writing and are considered refined language.

- (43) $\overset{\circ}{y}at$ $\tilde{t}iù$ $\overset{\prime}{h}o$ $\overline{h}on$
 one MEA good man
 a true man

- (44) $\overset{\circ}{y}at$ $\tilde{t}iù$ $\overline{d}aai$ $\overline{h}on$
 one MEA big man
 a physically strong man

The measure *dau* is another measure word which modifies *yáu* (bloke). It has the same register as *tiù* but it is less often used. Again the speaker uses the word to dehumanise the referent of the utterance because *dau* is normally the container measure for rice. A common vernacular expression is

- (45) $\overset{\circ}{y}at$ $\overset{\circ}{d}au$ \overline{faan}
 one MEA rice
 a large bowl [or any container of a similar function] of rice

In Cantonese, a very stupid person is called *faan-tung* ('rice-barrel'), who does not know how to do anything but eat (rice). Therefore by using *dau*, a measure word specific to rice and applying it to a person, the speaker shows his contempt for the referent.

At this point, we may be tempted to draw the conclusion that if a speaker uses a measure word which is normally used to modify an animal such as *tiu*, or an inanimate object, such as *dau*, to modify a person, he is displaying a negative attitude towards the referent. In fact, we cannot make such generalisations, as I will illustrate with another measure, *jek*.

The measure *jek* is an interesting individual measure for human beings. Normally it is used for animals, birds and insects. (See examples (7)–(10).) Sometimes Cantonese speakers use *jek* to compare a human being and an animal, as in

- (46) *Yau jek daai-jek lo.*
 have MEA big-sized man
 There is a physically strong man.

The focus is on the referent's physical size and the measure word relates his size to a large animal size. The sense is not derogatory. Sentence (47) shows another meaning:

- (47) *Kui jo do yat jek ye gam.*
 3SG works until one MEA creature like
 He/she is so busy with his/her work that he/she is just like a non-human creature.

The sentence probably has the same meaning as 'He works like a dog.' The measure word *jek* suggests the speaker's sympathy with the referent. Example (48) shows yet another meaning:

- (48) *Nei jek ye al*
 you MEA creature INT
 You little devil!

This is a term of gentle rebuke which indicates the intimate relationship between the speaker and the referent. When *jek* is contrasted with *tiu*, we find that describing a person as an animal does not necessarily entail a contemptuous attitude.

4. CONCLUSION

It can be seen that Cantonese measure words serve a variety of functions, from the simple semantic role of classifying human and non-human creatures according to their physical sizes and shapes to the more complex functions of expressing pragmatic roles and social interactions. We find that there is a much higher level of subtlety and flexibility in the use of measure words for human beings than for non-human creatures. This is probably true of measure words in other languages as well because human attributes are perceived as far more complex than those of non-human creatures, and there is far more fluidity in human relationships than in the relationship between humans and their environment. The users of a human language are constantly exploring ways of expressing these ever-changing relationships.

REFERENCES

- Adams, Karen L., 1986, Numeral classifiers in Austroasiatic. In Craig, ed. 1986:241–262.
1989, *Systems of numeral classification in the Mon-Khmer, Nicobarese and Aslian subfamilies of Austroasiatic*. PL, B–101.
- Becker, Alton L., 1975, A linguistic image of nature: the Burmese numerative classifier system. *International Journal of the Sociology of Language: Sociolinguistics in Southeast Asia* 5:109–121. The Hague: Mouton.
- Chao, Yuen Ren, 1968, *A grammar of spoken Chinese*. Berkeley, California: University of California Press.
- Craig, Colette, 1986, Jacalteco noun classifiers: a study in language and culture. In Craig, ed. 1986:263–293.
- Craig, Colette, ed., 1986, *Noun classes and categorization. Typological Studies in Language* 7. Amsterdam: John Benjamins.
- DeFrancis, John, 1984, *The Chinese language: fact and fantasy*. Honolulu: University of Hawaii Press.
- Denny, J. Peter, 1979, Semantic analysis of selected Japanese numeral classifiers for units. *Linguistics* 17:317–335.
- Dixon, R.M.W., 1986, Noun classes and noun classification in typological perspective. In Craig, ed. 1986:105–112.
- Erbaugh, Mary S., 1986, Taking stock: the development of Chinese noun classifiers historically and in young children. In Craig, ed. 1986:399–436.
- Juntanamalaga, Preecha, 1988, Social issues in Thai classifier usage. *Language Sciences* 10/2:313–330.
- Lakoff, George, 1986, Classifiers as a reflection of mind. In Craig, ed. 1986:13–51.
- Lam, Martha and Stanley Po, 1988, *Functional Cantonese Book One*. Ultimo: Melting Pot Press.
- NAATI (National Accreditation Authority for Translators and Interpreters, Australia), 1987, *NAATI tests: information (candidates' manual)*. Canberra.
- Norman, Jerry, 1988, *Chinese*. New York: Cambridge University Press.
- Ramsey, S. Robert, 1987, *The languages of China*. Princeton: Princeton University Press.
- Wong, S.L., 1941, 1995, *A Chinese syllabary pronounced according to the dialect of Canton*. Hong Kong: Chung Hwa Book Company.

THE STATUS OF THE WORD *hây* IN THAI

KITIMA INDRAMBARYA

1. INTRODUCTION¹

There has long been controversy about the word *hây* 'to give' in Thai: about how many homophonous words *hây* exist in Thai and to which grammatical categories they belong. This paper examines the status of the form *hây* based on distribution and meaning. Within the Lexicase version of dependency grammar, this analysis classifies the form *hây* into six homophonous variants: three ditransitive verbs, two causative verbs, and an adverb.

The form *hây* is most commonly a verb carrying the meaning 'to give'. However, there are several homophonous *hây*'s, as illustrated in the following example:

- (1) *Nốy bôk wáa mēē hây tǎy hây kuncēē dēē hây sǎm*
Noy say that mother cause Toy give key Dang cause repair
rót hây míchanán cà? tii hây.
car for otherwise will hit for
Noy said that mother ordered Toy, under the threat of being hit, to give the key
to Dang to fix the car for (mother).

The forms pronounced as *hây* in the example above differ in distribution and meaning. The purpose of this paper is to classify the Thai form *hây* into different categories based on their distribution and meaning. The syntactic framework used in this analysis is Lexicase, a version of dependency grammar developed by Stanley Starosta in the early 1970s. While the first part of this paper is an introduction, the second part is a review of previous analyses of *hây*. Sections 3–5 provide a reanalysis of the form *hây*: §3 discusses the ditransitive verbs *hây* 'to give'; §4 discusses the causative verbs *hây*; and §5 discusses the adverb *hây*. The last part of the paper presents a conclusion.

2. PREVIOUS ANALYSES

2.1 PANUPONG 1970

Vichin Panupong's work *Inter-sentence relations in modern conversational Thai* is a detailed structuralist analysis of spoken Thai based on the use of test frames. Panupong

¹ This article is a revision of Indrambarya (1990). I would like to thank Professor Stanley Starosta and Marybeth Clark for their valuable comments regarding previous versions of this article.

Dejthamrong defines ditransitive verbs in accordance with Panupong's (1970) ditransitive verb test frames. However, Dejthamrong notes that there are three types of ditransitive verbs: those that require a preposition, such as *sòŋ* 'to send', as in (7); those that do not allow a preposition, such as *tɔ̀ŋ* 'to answer' in (8); and those that may or may not have a preposition, such as *hây* 'to give' in (9).

- (7) a. **Lék sòŋ còtmăăy ɔrathay.*
 Lek send letter Orathai
 Lek sent the letter to Orathai.
- b. *Lék sòŋ còtmăăy kêe ɔrathay.*
 Lek send letter to Orathai
 Lek sent the letter to Orathai.
- (8) a. *Kháw tɔɔp khamthăam khɔɔ nîi khruu.*
 he answer question bunch this teacher
 He answered this question for the teacher. (lit. He answered this question the teacher.)
- b. **Kháw tɔɔp khamthăam khɔɔ nîi kêe khruu.*
 he answer question bunch this to teacher
 He answered this question for the teacher. (lit. He answered this question to the teacher.)
- (9) a. *Deeŋ hây náŋsĕ dèkdèk.*
 Dang give book children
 Dang gave the children some books.
- b. *Deeŋ hây náŋsĕ kêe dèkdèk.*
 Dang give book to children
 Dang gave some books to the children.

Verbs which require prepositions do not fit in Panupong's verb test frame. Dejthamrong's explanation for treating such verbs as ditransitive verbs is based on their overlapping distribution. They are included, according to Dejthamrong (1970:43), only because they may replace ditransitive verbs like *hây* 'to give' and *cêek* 'to distribute' when such verbs are followed by a preposition.

Hây in (10) and (11) is considered to be a preposition, since it fits in a preposition test frame, while *hây* in (12) is identified as a causative transitive verb followed by a clause.

- (10) *Deeŋ yím hây chán.*
 Dang smile give I
 Dang smiled at me.
- (11) *Deeŋ kwàat bāan hây mĕe.*
 Dang sweep house give mother
 Dang swept the house for her mother.
- (12) *Deeŋ hây nɔɔŋ kwàat bāan.*
 Dang cause younger.sibling sweep house
 Dang had his younger sibling sweep the house.

Dejthamrong considers *hây* in (13) and (15)–(18) to be instances of *hây* functioning as a "clause linker", which "introduces" a noun clause or an adverbial clause. For example, in (13a), *hây* is analysed as a clause linker, introducing a noun clause, because *hây chán maa* may be replaced by a noun *náŋsĕ* 'book', a direct object of the transitive verb *tɔŋkaan* 'to want', as in (13b) (Dejthamrong 1970:100).

- (13)a. *Dɛɛŋ ɿŋkaan hây chán maa.*
 Dang want I come
 Dang wants me to come.
- b. *Dɛɛŋ ɿŋkaan náŋsǎi.*
 Dang want book
 Dang wants a book.

When *hây* occurs with the ditransitive verb *bòk* ‘to tell, to order’, as in (14a), Dejthamrong treats the clause introduced by *hây* as a direct object of the verb because it may be replaced by the noun *khàaw* ‘news’, as shown in (14b), despite the meaning difference of the verb *bòk*. However, according to Dejthamrong, “unlike other ditransitive verbs”, which require a direct object to precede an indirect object, the indirect object *dɛɛŋ* ‘Dang’ precedes the direct object noun clause in (14a). To accommodate this fact, Dejthamrong (1970:101) stipulates that when the direct object is a noun clause, it appears after an indirect object.

- (14)a. *Kháw bòk dɛɛŋ hây klàp bâan.*
 he order Dang return home
 IO DO
 He ordered Dang to return home.
- b. *Kháw bòk khàaw dɛɛŋ.*
 he tell news Dang
 DO IO
 He told Dang the news.

In Dejthamrong’s analysis, the clause linker *hây* may introduce an adverbial clause modifying a verb, as in (15a) and (16a). *Hây* and the words that follow it constitute an adverbial clause because one could replace the clause with the adverb *yàaŋray* ‘how’, as in (15b), and *thammay* ‘why’, as in (16b) (Dejthamrong 1970:120–121).

- (15)a. *Chán cà? nɔɔn hây sabaay.*
 I will sleep comfortable
 I will sleep comfortably.
- b. *Chán cà? nɔɔn yàaŋray?*
 I will sleep how
 How will I sleep?
- (16)a. *Nít yók kâwǎi hây khèek nâŋ.*
 Nit lift chair guest sit
 Nit got a chair for the guest to sit on.
- b. *Nít yók kâwǎi thammay?*
 Nit lift chair why
 Why did Nit lift a chair?

Dejthamrong also considers the *hây* which can appear only at the beginning of a sentence to be a clause linker introducing an adverbial clause, as in (17a), because *hây* plus the words which accompany it may be replaced by *yaŋŋay yaŋŋay* ‘whatever happens’ in (17b) (Dejthamrong 1970:123).

- (17)a. *Hây fǒn tòk chán kǒw cà? pay roongrian.*
 rain fall I then will go school
 Even if it rains, I will still go to school.
- b. *Yanngayyanngay chán kǒw cà? pay roongrian.*
 whatever.happens I then will go school
 Whatever happens, I will still go to school.

Finally, Dejthamrong also sets up a category called postverb for the form *hây* which appears after a transitive verb. The postverbal *hây* functions as part of the nucleus, as illustrated in (18). However, she does not state to which syntactic category (e.g. noun, verb, adverb) this *hây* belongs.

- (18) *Mǎa kàt khâw hây.*
 dog bite he
 The dog bit him.

2.3 KULLAVANIJAYA 1974

Pranee Kullavanijaya (1974), working within an early version of the Lexicase framework, discusses the ditransitive verb *hây*, the causative verb *hây*, and the derived preposition *hây*. According to her analysis, a sentence like (19) is ambiguous. Read in one way, *hây* is considered to be a preposition, a benefactive case assigned to its following noun.

- (19) *Mĕĕ yĕp sĭa hây lûuk.*
 mother sew shirt for child
 +N +P +N
 +NM² +B +AC
 +AGT +BEN
 A mother sewed a dress for her child. (Kullavanijaya 1974:85)

Example (19) also has another reading, in which *hây* is a non-finite ditransitive verb. The sentence is interpreted as 'Mother sewed the dress and gave it to the child', as illustrated in (20).

- (20) *Mĕĕ yĕp sĭa hây lûuk.*
 mother sew shirt give child
 +N +V +N
 +NM -finite +AC
 +AGT +BEN

² The lexicase abbreviations for case markers used by Kullavanijaya (1974) are as follows:

Case Forms		Case Relations	
AC	Accusative	AGT	Agent
B	Benefactive	BEN	Benefactive
C	Comitative	DAT	Dative
NM	Nominative		

- (26) *Surii khâprôt hây suda.*
 Suri drive for Suda
 Suri drives for Suda. (Thepkarnchana 1986:200)

Coverbs are “semantically depleted”. According to Thepkarnchana (1986:201), *hây* in (26) does not lose its semantic properties; it only loses its literal sense – to give something to Suda. However, it maintains the metaphorical interpretation, namely a favour that *Suri* bestows on *Suda*. From this interpretation, Thepkarnchana concludes that *hây* in this construction is a verb. However, her claim seems to contradict itself. If a word which loses its literal sense is not a coverb, what is semantically depleted?

Thepkarnchana advances two syntactic arguments for claiming that all words which have corresponding verbs are verbs. First, these words have the potential to occur as full verbs in isolation, while maintaining the same meaning. Thus, they can be analysed as verbs in other contexts as well. For example, the word *khâam* ‘to cross’ in (27a) may appear as a full verb in (27b) and is thus regarded as a verb rather than a preposition.

- (27)a. *Kháw dæn khâam saphaan pay.*
 he walk cross bridge go
 He walked across the bridge.
- b. *Kháw khâam saphaan pay.*
 he cross bridge go
 He crossed the bridge. (Thepkarnchana 1986:205–206)

However, applying the same test to (26), we find that *hây* cannot appear as a free verb without changing the meaning from ‘for’ to ‘to give’, as shown in (28). Hence, by Thepkarnchana’s own criterion, *hây* in this instance is not a verb.

- (28) *Surii hây suda.*
 Suri for Suda
 *Suri for Suda.
 Suri gave Suda (something).

Second, if a word can be negated, it is a verb. Thepkarnchana exemplifies this claim with *khâam* ‘to cross’.

- (29) *Kháw kradòot mây khâam rúa.*
 he jump not cross fence
 He jumped, but failed to hurdle the fence. (Thepkarnchana 1986:206)

Since *khâam* ‘to cross’ in (30) may be negated, Thepkarnchana concludes that all words in Thai which have corresponding verbs have verbal status synchronically. However, Thepkarnchana has neglected to note that some adverbs, such as *bây* ‘often’, may also be negated, as shown in (30b). *Bây* is clearly not a verb, since it never appears as a free verb in Thai, as illustrated in (30c). It is evident that the negation test cannot reliably distinguish a verb from an adverb when the word in question occurs after another verb.

- (30)a. *Kháw maa thîi nîi bây.*
 he come at here often
 He comes here often.

- b. *Kháw maa thîi nîi mây bôy.*
 he come at here not often
 He does not come here often.
- c. **Kháw bôy.*
 he often
 *He often.

3. DITRANSITIVE VERBS *hây*

Hây meaning 'to give' always functions as a verb. I claim that there are three ditransitive verbs *hây* 'to give'. *Hây*₁ requires two bare noun phrases as its complements, as in (31). *Hây*₂ requires, as its complements, a bare noun phrase followed by a prepositional phrase, as in (32). *Hây*₃ differs from *hây*₁ only in that it requires an additional verb complement and forbids a PP, as shown in (33).

- (31) *Lék hây₁ còtmăăy dɛɛŋ.*
 Lek give letter Dang
 Lek gave Dang a letter.
- (32) *Lék hây₂ còtmăăy kêɛ dɛɛŋ.*
 Lek give letter to Dang
 Lek gave a letter to Dang.
- (33) *Chán hây₃ náŋsǎi dèkdèk ʔàan.*
 I give book children read
 I gave a book to the children to read.

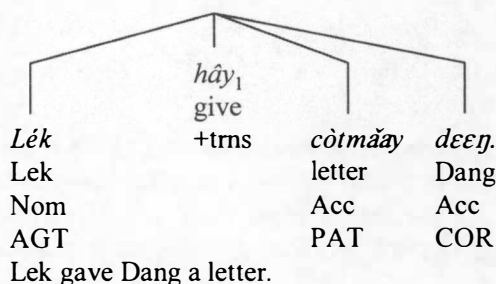
In Thai, the nominal case forms, that is the case markings without a preposition, are found to be limited to the Nominative, Accusative, and Locative case forms. The Nominative marks a Patient (PAT) of an intransitive finite verb and an Agent (AGT) of a transitive finite verb.³ The Locative marks the locational noun of a locational verb as Locus case relation (LOC). Patient of transitive verbs and Correspondent actants are realised as the Accusative case form. The Locus actant of a non-locational verb and a Means actant, on the other hand, always occur as an immediate dependent of a preposition and carry the Accusative case form assigned by the regent preposition (Indrambarya 1994:62). The Patient Centrality Hypothesis, together with the One per Sent constraint, ensures that every verb must have a Patient in its case frame and that a clause may carry only one Patient (Starosta 1988:128,138). Since the ditransitive verb *hây*₁ in (34) has two accusative nouns, only one of them will be assigned as Patient. The other must be a Correspondent actant. The question is which noun is assigned Patient and which Correspondent.

In a passive construction, an object (Patient) of a transitive verb is omitted and is interpreted as the Patient of the verb *thiuk* or *doon* (see P2P Control Rule in Appendix IV). Since only a transitive verb may be passivised in Thai, a passive construction may serve as

³ Locational verbs are verbs which expect a dependent which is lexically marked as [+lctn]. These verbs are, for example, *pay* 'to go' and *yuu* 'to stay' (Indrambarya 1994:53).

a test for the Patient case relation. Unfortunately, this process is not very productive in Thai and is subject to certain semantic constraints, so that neither of the two objects of *hây*₁ nor those of other ditransitive verbs may occur as Patient of *thùuk*, as shown in (34b) and (34c). In this paper, I will assume that the semantically transferred object of the ditransitive verb *hây*₁, such as *còtmăăy* 'letter', is assigned PAT in the same way that the bare transferred object of the ditransitive verb *hây*₂, which is the only accusative actant of the verb, is assigned PAT. Hence, in (34), *còtmăăy* 'letter' is PAT and *dɛɛŋ* 'Dang' is COR.

(34)a.



- b. **Còtmăăy thùuk lék hây₁ dɛɛŋ.*
 letter undergo Lek give Dang
 A letter was given to Dang by Lek.

- c. **Dɛɛŋ thùuk lék hây₁ còtmăăy.*
 Dang undergo Lek give letter
 Dang was given a letter by Lek.

Before arguing whether a form belongs to two separate lexical entries, one first needs to distinguish a complement, an argument that subcategorises a verb, from an adjunct, which by general rule may optionally occur with that verb. The distinction is tested here in two ways: by the Question Pull test and by the head substitution criterion.

Since Thai is a 'pro-drop language' (Cole 1987:606–607), that is a language which makes significant use of zero anaphora, any of the three arguments of *hây*₁ may be omitted contextually. When the arguments of verbs are omitted, the Question Pull test (Hasagawa 1988) may help to determine whether a verb is transitive or ditransitive. According to Hasagawa, given an utterance consisting of a verb with few or no other accompanying words, an addressee is expected to ask What-Who-questions about obligatory arguments, that is, complements, which are not clear in a conversation. Questions introduced by 'why', 'where', and 'when', which extract adjuncts, will be asked only after all required arguments are retrieved. Consider example (35):

- (35) Speaker A: *lék hây còtmăăy*
 Lek give letter
- Speaker B: *hây khray*
 give who
- Speaker A: *hây dɛɛŋ*
 give Dang

Because speaker A gave insufficient information about the word *hây* 'to give' to speaker B, speaker B is expected to inquire as to whom *lék* 'Lek' gave a letter. This shows that *hây* in (35) is looking for three arguments and must be interpreted as the ditransitive verb *hây*₁.

The head substitution test is another criterion for complement and adjunct distinction. Heads may differ in their ability to cooccur with a particular complement type, while an adjunct may occur freely with any head, subject to pragmatic considerations. Taking *kêê dɛɛŋ* 'to Dang' in (36) as an example, when the head *hây₂* is replaced with one or another verb denoting a transfer meaning, *kêê dɛɛŋ* 'to Dang', if it is a complement, should not be able to occur with such verbs, subject to pragmatic constraints. Example (36) shows that this is in fact the case. Since *tɔɔp* 'to answer' may not appear with *kêê dɛɛŋ*, I conclude that *kêê dɛɛŋ* in (36) is a complement of *hây₂*.

- (36) *Lék hây₂/*tɔɔp còtmăăy kêê dɛɛŋ.*
 Lek give/answer letter to Dang
 Lek gave/*answered the letters to Dang.

Hây₂ is a locative ditransitive verb requiring three arguments. In (37), the subject *khruu* is AGT, and *hây₂* takes only one accusative actant. If a verb expects only one object, that object must be assigned Patient according to the Patient Centrality Hypothesis. Hence *náŋsĥĩ* must be PAT. *Lék* 'Lek' is assigned the LOC case relation since it is the accusative actant of the locative preposition *kêê* 'to'.

- (37)
-
- | | | | | |
|--------------|--------------------------------|---------------|------------|-------------|
| <i>Khruu</i> | <i>hây₂</i>
give | <i>náŋsĥĩ</i> | <i>kêê</i> | |
| teacher | +trns
+lctn | book | to | |
| Nom | | Acc | +P | <i>lêk.</i> |
| AGT | | PAT | +goal | Lek |
| | | | | Acc |
| | | | | LOC |

The teacher gave a book to Lek.

The head substitution test below illustrates that *ʔaan* 'to read' in example (33) is a complement of *hây₃*. Only *hây* but not *cêêk* 'to distribute' may occur with the embedded clause containing *ʔaan*, as shown in (38).

- (38) *Chán hây₃/*cêêk náŋsĥĩ dèkdèk ʔaan.*
 I give/distribute book children read
 I gave/distributed books to the children to read.

In Lexicase grammar, a form is considered to be two distinct lexical entries if it occurs in two distinct GRAMMATICALLY SIGNIFICANT ENVIRONMENTS A and B and there is a word X which may appear in pattern A but not in B and/or there is a word Y which may appear in B but not in A (Starosta 1988:98). Consider (39) and (40):

- (39) a. *Nít tɔɔm námman rôt khan nii lééw.*
 Nit fill oil car vehicle this already
 Nit has already filled this car with gas.
- b. **Nít tɔɔm námman kêê rôt khan nii lééw.*
 Nit fill oil to car vehicle this already
 Nit has already filled the gas into this car.

- (40)a. **Kulayaa mɔ̌ɔp còtmǎ̌y chabàp nán khruu.*
 Kulaya deliver letter issue that teacher
 Kulaya delivered that letter to the teacher.
- b. *Kulayaa mɔ̌ɔp còtmǎ̌y chabàp nán kɛ̌ khruu.*
 Kulaya deliver letter issue that to teacher
 Kulaya delivered that letter to the teacher.

Tɔ̌əm 'to fill' in (39) and *mɔ̌ɔp* 'to deliver' in (40) can each appear in only one of the two constructions. *Tɔ̌əm* 'to fill' may appear only in (39a) but not in (39b), while *mɔ̌ɔp* 'to deliver' may appear in (40b) but not in (40a). The distribution shows that the two patterns exemplify two grammatically significant environments.

By this criterion, if a single form appears in both environments, it must belong to two separate verb classes and constitute two distinct lexical entries, since if two forms differ in either pronunciation, meaning, or distribution, they must be learned and stored separately. Thus, from the observation that the form *hây* can appear in both of the frames (a) and (b), as shown in examples (31) and (32), we can conclude that they are two distinct lexical items. *Hây*₁ belongs to the same verb class as *tɔ̌əm* 'to fill' in (39a), and *hây*₂ belongs to the same verb class as *mɔ̌ɔp* 'to deliver' in (40b).

The grammatically significant environments criterion also shows that *hây*₁ in (31) and *hây*₃ in (33) belong to two different classes. Consider (41):

- (41)a. *Chán cè̌ek₁ náŋsǎ̌ dèkdèk.*
 I distribute book children
 I distributed the books to the children.
- b. *Chán cè̌ek₂ náŋsǎ̌ kɛ̌ dèkdèk.*
 I distribute book to children
 I distributed the books to the children.
- c. **Chán cè̌ek náŋsǎ̌ dèkdèk ʔàan.*
 I distribute book children read
 I distributed the books to the children to read.

Again, while the form *hây* may appear in both constructions (31) and (33), the form *cè̌ek* 'to distribute' can be used in (41a) but not in (41c). Unfortunately, an example of a ditransitive verb other than *hây* which could occur in (41c) but not in (41a) has not been found. Nevertheless, the distribution exemplified in (41a) and (41c) is sufficient to show that the forms *hây* in the two patterns are two different lexical entries.

The words A and B are related by the synchronic derivational rule if they are distinct lexical items which are etymologically related and if (A:B) is a member of an analogical set {X:Y} which contains other pairs of lexical items related in the same way (Starosta 1988:63). Example (41) demonstrates that *hây*₁ and *hây*₂ are derivationally related, since there is the form *cè̌ek* which may appear in both the (a) and (b) patterns of (41). However, since no other ditransitive verbs may appear in the same environment as *hây*₃, there is no derivational relationship between *hây*₁ and *hây*₃ or *hây*₂ and *hây*₃. The derivational rule which relates *hây*₁ to *hây*₂ may be formulated as follows (see abbreviations in Appendix I):

(42)	<table> <tr><td>+crsp</td></tr> <tr><td>+trns</td></tr> <tr><td>-lctn</td></tr> <tr><td>-xtns</td></tr> <tr><td>n[+COR]</td></tr> </table>	+crsp	+trns	-lctn	-xtns	n[+COR]	:	<table> <tr><td>+goal</td></tr> <tr><td>+lctn</td></tr> <tr><td>+trns</td></tr> <tr><td>-crsp</td></tr> <tr><td>n[+LOC]</td></tr> </table>	+goal	+lctn	+trns	-crsp	n[+LOC]
+crsp													
+trns													
-lctn													
-xtns													
n[+COR]													
+goal													
+lctn													
+trns													
-crsp													
n[+LOC]													

The derivational rule in (42) says that the ditransitive verb *hây₁* is related to the locative ditransitive verb *hây₂*. The contextual features n[+COR] in the left column and the contextual features n[+LOC] in the right column illustrate that the Correspondent actant in *hây₁* corresponds to the Locus actant in *hây₂*.

I consider the embedded verb *ʔàan* 'to read' in (33) (repeated as (43a) below) to be a transitive verb, the object of which is missing. This conclusion is based on the fact that *ʔàan* can be followed by an object which refers to part of the higher object, as shown in (43b).

- (43) a. *Chán hây₃ náŋsǎi dèkdèk ʔàan.*
 I give book children read
 AGT +trns PAT COR +trns
 I gave a book to the children to read.
- b. *Chán hây₃ náŋsǎi dèkdèk ʔàan náapòk.*
 I give book children read cover
 AGT +trns PAT COR +trns PAT
 I gave a book to the children to read the cover.

The reason that I do not consider *ʔàan* in (43a) and (43b) to be two separate lexical items, an intransitive and transitive pair, is that I cannot find verbs which may occur in one of the two constructions but not the other. By the grammatically significant environments criterion, *ʔàan* in the two examples must therefore be regarded as a single lexical item.

In (43a) the missing object of the transitive verb *ʔàan* 'to read' is interpreted as *náŋsǎi* 'book', which is the higher Patient (PAT). This relationship between the missing lower object and the higher object is stated informally in Pagotto's account of English 'Tough' movement (Pagotto 1985:42). In Thai, I find that a similar relationship holds in verbs which carry the feature [+cntn] (continuing). The feature [+cntn] shows the continuation of actions between the regent verb and its dependent verb. I will formulate this relationship in terms of Lexicase control chaining rules, rules which state the relationship between words in different domains. Furthermore, the relationship between a lower direct object in (43b), which refers to part of the higher object, and the higher direct object could be accounted for by a similar rule which will not be formulated in this paper. The coreference between a missing Patient of an infinitival verb complement and the Patient of a regent continuing verb in example (43a) is accounted for by the Patient-to-Patient Control Chaining Rule,

4 Domain: direct and indirect relationship between a lexical head and its dependent. Y is in the domain of X if X is the regent of Y.

hereafter called the P2P Control Rule (Indrambarya 1994:299–301), as shown in (44).⁵ This rule applies only to a subset of verbs taking infinitival complement clauses.

(44)

- a. $\left[\begin{array}{c} -\text{fint} \\ ?[+\text{PAT}] \\ \text{nndex} \end{array} \right] \rightarrow [m[+\text{PAT}]] \setminus \left[\begin{array}{c} +\text{cntn} \\ m[+\text{PAT}] \\ n[-\text{fint}] \end{array} \right]$
- b. The actor of the non-finite verb is interpreted as the closest available noun-headed dependent of a regent verb to the left of the embedded verb. (Indrambarya 1994:300)

As shown in (44), the P2P rule consists of two parts, one looking for a missing Patient of the lower clause in (44a) and another looking for an actor of the lower clause in (44b), only the first of which is formulated. A non-finite verb with *nndex* is looking for a Patient shown by $[?[\text{PAT}]]$.⁶ The non-finite verb will find its Patient by copying the index *m* from $[m[+\text{PAT}]]$ of the regent. The relationship between the non-finite verb and its regent is shown by the feature $[n[-\text{fint}]]$ on the regent. That is, a regent is looking for a non-finite verb with an index *n*. The feature $[+\text{cntn}]$ restricts this rule to apply only to continuing verbs and shows the continuation of actions between the regent verb and its dependent verbs (Indrambarya 1994:299–304). Moreover, the Redundancy Rule 1 (RR-1) states that the feature $[+\text{cntn}]$ expects a non-finite transitive verb as its dependent, indicated by question mark in front of the features. When the requirement is fulfilled, each question mark is replaced by an index number (see Appendix II for members of continuing verbs in Thai).

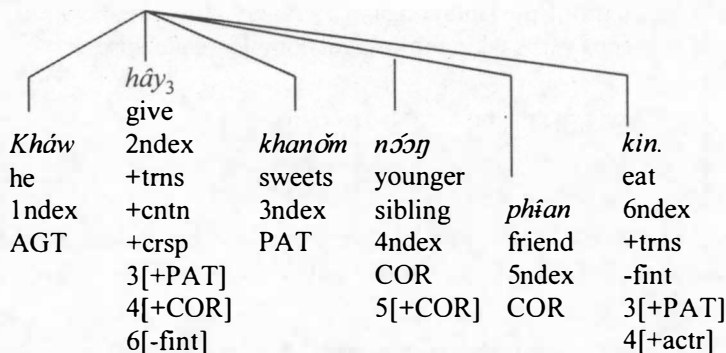
$$\text{RR-1} \quad [+ \text{cntn}] \rightarrow \left[\begin{array}{c} +\text{V} \\ ?[-\text{fint}] \\ ?[+\text{trns}] \end{array} \right]$$

The missing actor of the infinitival complement verb, on the other hand, is interpreted as the closest available noun-head dependent of a regent verb to the left of the embedded verb (see Pagotto 1985:44–45). The term ‘available nominal’ refers to a noun which is free in its governing category following Principle B of the Binding theory (Chomsky 1981:181). A governing category is defined here as the domain of the verb containing the missing Patient. In other words, to satisfy the binding requirement, an implied actor to the left of the embedded verb must not be coreferential with the missing Patient. As an example, consider (45).

⁵ This rule is a revision of the Patient-to-Patient Control Rule proposed in Indrambarya (1990:43).

⁶ The first letter *n* in *nndex* denotes the number for the index which specifies a given word’s position in a sentence.

(45)



He gave his friend's younger sibling sweets to eat.

According to the P2P rule in (44), the missing Patient of the verb *kin* 'to eat' is coreferential with the higher Patient *khanǒm* 'sweets'. This implied coreferentiality is shown by [3[+PAT]] on the lexical matrix of *kin*. To search for the actor of the embedded verb *kin*, one looks to the left of *kin* to find two noun-headed dependents of the regent verb *háy₃*, namely *khanǒm* 'sweets' and *nǒŋ phīan* 'friend's younger sibling'. If *khanǒm* were chosen as an implied actor for *kin*, the Binding Principle would be violated. Since *khanǒm* is already selected as a coreferential Patient, it is not an available noun. Hence, *nǒŋ phīan* 'friend's younger sibling' is the implied actor for *kin* because it is the available noun closest to the embedded verb *kin*. This fact is shown by the implied feature 4[+actr] on the lexical matrix of *kin*.

Similarly, when the ditransitive verb *háy₃* appears in an embedded clause, as in (46), the missing Patient of *háy₃* is chained to the Patient *khanǒm* 'sweets' of the higher verb *s#i* 'to buy', while *dεεŋ* 'Dang', the first available noun phrase to the left of *háy₃*, is the actor. However, if the closer nominal *khanǒm* were the actor, *háy₃* would have 3[+PAT] as both actor and Patient, violating the Binding Principle. Likewise, the Patient of the lower verb *kin* 'to eat' is interpreted as *khanǒm* 'sweets' by the feature 3[+PAT] in the same way as is the object of *háy₃*. The actor of *kin* is the closest available nominal, *dèkdèk*.

(46)	<i>Dεεŋ</i>	<i>s#i₂</i>	<i>khanǒm</i>	<i>háy₃</i>	<i>dèkdèk</i>	<i>kin₁</i>
	Dang	buy	sweets	give	children	eat
	1ndex	2ndex	3ndex	4ndex	5ndex	6ndex
	AGT	+trns	PAT	+trns	COR	+trns
		+cntn		+cntn		-fint
		+crsp		+crsp		3[+PAT]
		3[+PAT]		-fint		5[+actr]
		4[-fint]		3[+PAT]		
				5[+COR]		
				1[+actr]		
				6[-fint]		

Dang bought sweets to give to the children to eat.

The embedded clause containing *háy₃* is considered to be a complement. In applying the head substitution test in (47), the heads *mǒŋ* 'to look' and *duu* 'to search' cannot occur with the embedded *háy* clause, while *hǎa* 'to search' can.

- (47) *Dɛɛŋ *duul/*mɔɔŋ/hǎa khanǒm hây₃ dèk kin₁.*
 Dang search/look/search sweets give child eat
 Dang looked for sweets to give to the child to eat.

The ditransitive verbs *hây₁* and *hây₂*, as in (48) and (49), may appear in an embedded clause with a missing Patient, and undergo the P2P rule in the same way as *hây₃* does. The index 3[+PAT] on *hây₁* and *hây₂* comes from the index 3[+PAT] of the higher verb *s#₂* 'to buy', showing that the Patients of both verbs are coreferential. *Dɛɛŋ* 'Dang', the first available NP to the left of the *hây* clause, is the actor of *hây₁* as well as *hây₂*.

- (48) *Dɛɛŋ s#₂ náŋsǎi hây₁ lék.*
 Dang buy book give Lek
 1ndex 2ndex 3ndex 4ndex 5ndex
 AGT +trns PAT +trns COR
 +cntn +cntn
 3[+PAT] +crsp
 1[+actr] -fint
 4[-fint] 3[+PAT]
 1[+actr]
 Dang bought a book to give to Lek.
- (49) *Dɛɛŋ s#₂ náŋsǎi hây₂ kèɛ lék.*
 Dang buy book give to Lek
 1ndex 2ndex 3ndex 4ndex 5ndex 6ndex
 AGT +trns PAT +trns +P LOC
 +cntn +cntn +goal
 3[+PAT] +lctn
 1[+actr] -fint
 4[-fint] 3[+PAT]
 1[+actr]
 Dang bought the book to give to Lek.

The forms *hây* in (48) and (49) are considered to be verbs (as opposed to the adverb *hây₆*, to be discussed in §5 for the following reasons. First, as shown in (50a) and (50b), the forms *hây* in examples (48) and (49) above may appear as finite verbs independently, preserving the meaning of *hây₁* and *hây₂*, respectively, and omit an understood object.

- (50)a. *Dɛɛŋ hây₁ lék.*
 Dang give Lek
 Dang gave (it) to Lek.
- b. *Dɛɛŋ hây₂ kèɛ lék.*
 Dang give to Lek
 Dang gave (it) to Lek.

Second, the auxiliary adverb *cà?* 'will' may precede a verb, but not an adverb or a preposition (Noss 1964, Indrambarya 1994). For example:

- (51)a. *Chán cà? pay duu nǎŋ phrûŋnîi.*
 I will go look movie tomorrow
 I will go to see a movie tomorrow.

- b. *Chán tâŋcay cà? pay duu năŋ phrúgnii.*
 I intend will go look movie tomorrow
 I intend to go to see a movie tomorrow.
- c. **Chán pay duu năŋ cà? bôy.*
 I go look movie will often
 I go to see movies often.
- d. **Chán pay duu năŋ cà? kâp phian.*
 I go look movie will with friend
 I go to see movies with a friend.

*Hây*₁ and *hây*₂ in an embedded clause may be preceded by *cà?*, which shows them to be verbs. Another piece of supporting evidence is based on a semantic test, namely the number of actions implied by the sentence. According to Li and Thompson (1973:267), if a sentence contains more than one action, there must be more than one verb. Since (48) and (49) contain two actions, to buy a book and to give it to *lék*, the evidence suggests that each sentence has two verbs, *sii* 'to buy' and *hây* 'to give'.

- (52)a. *Deeŋ sŋ₂ náŋsŋ^h cà? hây₁ lék.*
 Dang buy book will give Lek
 Dang bought a book to give to Lek.
- b. *Deeŋ sŋ₂ náŋsŋ^h cà? hây₂ kēē lék.*
 Dang buy book will give to Lek
 Dang bought a book to give to Lek.

4. CAUSATIVE VERBS *hây*

This section discusses two causative verbs *hây*: the personal causative verb *hây*₄ and the impersonal causative verb *hây*₅.

4.1 PERSONAL CAUSATIVE VERB *hây*₄

The causative verb *hây*₄, which carries the meaning of 'to order, to cause or to allow', is a non-verbal causative transitive verb which requires a sentential complement. Consider (53a) and (53b):

- (53)a. *Nit hây lék càt dōōkmáay.*
 Nit cause Lek arrange flower
 Nit had Lek arrange the flowers.
- b. **Nit khōōsōŋ lék càt dōōkmáay.*
 Nit plead Lek arrange flower
 Nit pleaded with Lek to arrange the flowers.

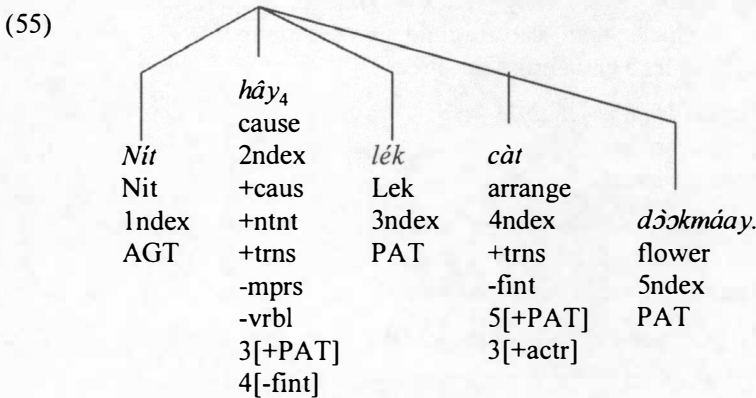
⁷ Two major types of causative verbs are non-verbal and verbal causative verbs. *Hây*₄ and *thamhây* are instances of non-verbal causative verbs. Verbal causative verbs are, for example, *khōōsōŋ*₂ and *sāŋ*₂ (Indrambarya 1994:279).

The head substitution test tells us that the embedded clause is a complement, because only *hây* but not *khôj-jôj* 'to plead' may appear with the embedded clause.

The verb *thiuk* 'to undergo' marks passive construction in Thai. Its subject is coreferential with the missing Patient object of the embedded transitive. Only transitive verbs, though not all transitive verbs, may occur in the *thiuk* passive construction. Since *hây₄* in (53) may occur in the *thiuk* passive construction, as illustrated in (54), *lêk* 'Lek' as subject with *hây₄* is analysed as transitive verb in this study. This evidence eliminates the possibility that *hây₄* might be an intransitive verb taking a finite complement. In the Lexicase framework, a finite verb, by definition, is a verb which allows a subject.

- (54) *Lêk thiuk mêe hây₄ càt dâjkmáay.*
 Lek undergo mother cause arrange flower
 Lek was ordered by her mother to arrange the flowers.

The tree structure for (53a) is shown in (55). The personal causative transitive verb *hây₄* is marked in its lexical matrix with the feature [+caus,+ntnt,+trns,-mprs,-vrbl] (causative, intentional, transitive, non-impersonal, and non-verbal) to be distinguished from other causative verbs in Thai (Indrambarya 1994:281). According to the Regular Actor Control Rule for infinitival complements (Starosta 1988:133; see Appendix IV), *lêk* 'Lek' is a Patient of the higher verb *hây* as well as the actor of the lower clause.



Nit had Lek arrange the flowers.

As with other Thai transitive verbs, the Accusative Patient of the causative transitive verb *hây* may be omitted. In this case, the Patient of the matrix verb is contextually bound and may be recovered by the index *x* in the lexicase external linking rule. The simplified version of the external linking rule is shown in (56). On the other hand, the actor of the lower non-finite complement verb is accounted for by the Regular Actor Control Rule. In other words, there are two types of missing constituents: one is a null anaphor which can be replaced by an overt noun and recovered by a widely applicable external linking rule; the other cannot be replaced by an overt noun and must be accounted for by a control chaining rule.

- (56) $\left| \begin{array}{l} +trns \\ ?[+PAT] \end{array} \right| \rightarrow [x[+PAT]]$

- (57) *Mêe hây₄ càt dâjkmáay.*
 mother cause arrange flower
 Mother had someone arrange the flowers.

In (57), the missing object of a transitive verb *hây*₄ is not grammatically recoverable at the sentence level and so is given an arbitrary index x, as [x[+PAT]]. The index x is also assigned to the actor of the lower clause yielding [x[actr]] in (57), according to the Regular Actor Control Rule for infinitival complements.

The causative verb *hây*₄ may appear in an embedded clause as a complement of a manipulative verb. One reason for identifying this form *hây* as a verb is that it may also appear as a main verb with the same syntactic frame and semantic reading, as shown in (58b) and (59b).

- (58)a. *Phố sàŋ₂ manít hây₄ klàp bân.*
 father order Manít cause return house
 Father ordered Manít to return home.
- b. *Phố hây₄ manít klàp bân.*
 father cause Manít return house
 Father had Manít return home.
- c. **Phố sàŋ₂ manít phố hây₄ klàp bân.*
 father order Manít father cause return house
 Father ordered Manít to return home.
- (59)a. *Kháw yùt rôt khan nán hây₄ dèk khâam thanố.*
 he stop car vehicle that cause child cross street
 He stopped that car to let a child cross the street.
- b. *Kháw hây₄ dèk khâam thanố.*
 he cause child cross street
 He let a child cross the street.
- c. **Kháw yùt rôt khan nán kháw hây₄ dèk khâam thanố.*
 he stop car vehicle that he cause child cross street
 He stopped that car to let a child cross the street.

The unacceptability of (58c) and (59c) when a subject is inserted shows that the embedded *hây*₄ in (58a) and (59a) is non-finite.

I will apply the head substitution test to (58a) and (59a) to see whether the embedded clause introduced by *hây*₄ is a complement. Consider (60) and (61):

- (60) *Phố *hây₄/bốc/*tham manít hây₄ klàp bân.*
 father cause/order/make Manít cause return house
 Father caused/ordered/made Manít (to) return home.
- (61) *Kháw yùt/cốt/hâam rôt hây₄ dèk khâam thanố.*
 he stop/park/halt car cause child cross street
 He stopped/parked/physically halted the car to let a child cross the street.

The unacceptability of *hây*₄ and *tham* in (60) shows that the embedded clause containing *hây*₄ is a complement. On the other hand, the embedded clause in (61) is an adjunct because all of the three verbs *yùt*, *cốt*, and *hâam* can occur with the embedded clause.

The following subsection discusses the characteristics of *hây*₄ in an embedded clause. Verbs which require the presence of the causative *hây*₄ as their verb complement are referred to as manipulative verbs in this study. Semantically, manipulative verbs may be

divided into two subclasses: manipulative verbal causative verbs and manipulative resultative verbs.

4.1.1 PERSONAL CAUSATIVE VERB *hây* EMBEDDED UNDER MANIPULATIVE VERBAL CAUSATIVE VERBS

The personal causative transitive verb *hây*₄ may be embedded under manipulative verbal causative verbs. Manipulative verbal causative verbs are verbs which show the intention of speakers to cause a Patient to undergo certain processes or actions, and which require the presence of the *hây*₄ clause as their dependent. There are three classes of manipulative verbal causative verbs: non-correspondent manipulative verbal causative intransitive verbs, Correspondent manipulative verbal causative intransitive verbs, and manipulative verbal causative transitive verbs.

4.1.1.1 NON-CORRESPONDENT MANIPULATIVE VERBAL CAUSATIVE INTRANSITIVE VERBS

Non-correspondent manipulative verbal causative intransitive verbs are intransitive verbs which require only a Patient in their case frame. They include *anúŷyâat* 'to allow', *wăŷ* 'to hope', *yàak* 'to want', *khăŋ* 'to ask', *khăŋŋŋ*₁ 'to plead', *sàŋ*₁ 'to order', and so forth. In example (62), the actor of *hây*₄ is interpreted as the upper Patient *dēēŋ* 'Dang' in accordance with the Regular Control Rule for infinitival complements (P2a).

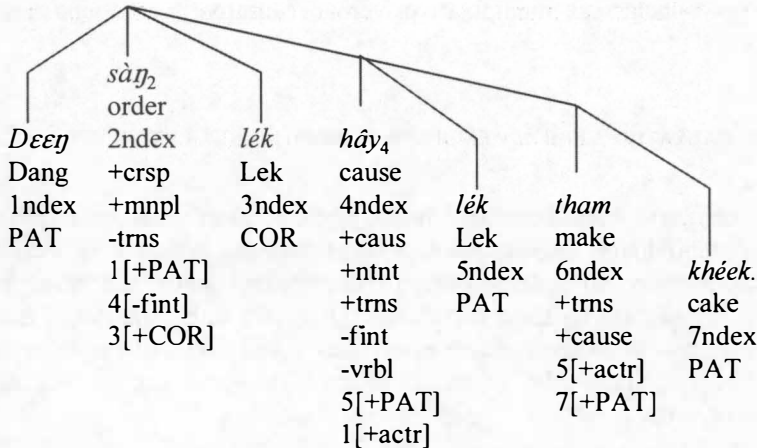
(62)	<i>Dēēŋ</i>	<i>khăŋŋŋ</i> ₁	<i>hây</i> ₄	<i>lék</i>	<i>tham</i>	<i>khéek</i> .
	Dang	plead	cause	Lek	make	cake
	1index	2ndex	3ndex	4ndex	5ndex	6ndex
	PAT	+caus	+caus	PAT	+trns	PAT
		+mnpl	+ntnt		-fint	
		+vrbl	-fint		4[+actr]	
		-trns	-mprs		6[+PAT]	
		1[+PAT]	+trns			
		1[+actr]	-vrbl			
		3[-fint]	4[+PAT]			
			1[+actr]			
			5[-fint]			

Dang pleaded with Lek to make a cake.

4.1.1.2. CORRESPONDENT MANIPULATIVE VERBAL CAUSATIVE INTRANSITIVE VERBS

Correspondent manipulative verbal causative intransitive verbs require both Patient and Correspondent case relations in their case frames. They are, for example, *sàŋ*₂ 'to order', *khăŋŋŋ*₂ 'to plead', *waan*₂ 'to ask', *băŋk*₂ 'to order', and so forth. The object of the embedded verb *hây* is obligatory for these verbs. These verbs have manipulative verbal causative transitive verb homophones. Again, by the P2a Control Rule, *hây* finds its actor from the higher Patient *lék*.

(63)a.



Dang ordered Lek to make a cake. (lit. Dang ordered Lek to cause Lek to make a cake.)

(63)b. **Lék nà? dəɛŋ sàŋ₂ hây lék tham khéek.*

Lek TOP Dang order cause Lek make cake

As for Lek, Dang ordered Lek to make a cake. (lit. Lek, Dang ordered to cause Lek to make a cake.)

(63)c. **Lék thiuk₅ dəɛŋ sàŋ₂ hây lék tham khéek.*

Lek undergo Dang order cause Lek make cake

Lek was ordered by Dang to make a cake.

(63)d. **Dəɛŋ sàŋ lék.*

Dang order Lek

Dang pleaded with Lek.

The verb *sàŋ₂* is analysed as a correspondent intransitive verb rather than a transitive one, because the noun phrase *lék* may not be topicalised, as shown in (63b), in contrast to an Accusative Patient of the homophonous transitive verb *sàŋ₃*, as shown in (64b), in §4.1.1.3.

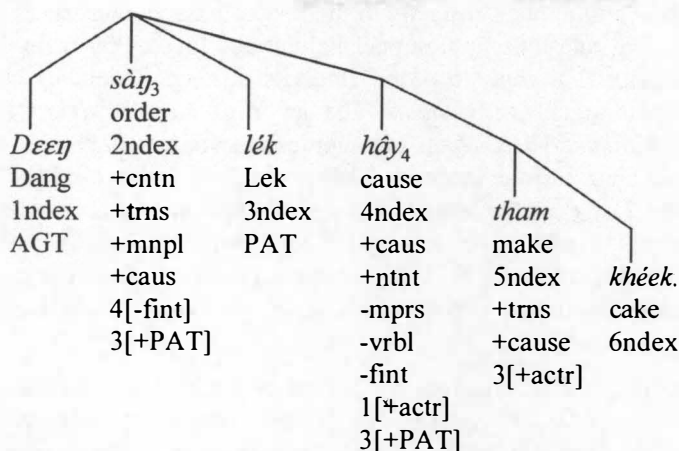
Supporting evidence for treating *sàŋ₂* as a correspondent intransitive verb is the fact that the verb *sàŋ₂* may not occur as the dependent of a continuing verb, such as *thiuk₅*, which are interpreted by the P2P rule, as shown in (63c) (as opposed to *sàŋ₃* in (64c) in the next subsection).

The embedded clause headed by *hây* is considered a complement because it is obligatory, as shown in the unacceptable (63d) without context.

4.1.1.3 MANIPULATIVE VERBAL CAUSATIVE TRANSITIVE VERBS

Manipulative verbal causative transitive verbs require both Agent and Patient in their case frames. They include *sàŋ₃* 'to order', *khǎŋǎŋ₃* 'to plead', *waan₃* 'to ask', *bǎŋk₃* 'to order', and so forth. The evidence for treating these verbs as transitive verbs stems from the fact that the missing object of *hây* is recovered by the P2P Control Rule, that the noun phrase *lék* can be topicalised, as in (64b), and that there is a related *thiuk* passive construction, as in (64c).

(64)a.



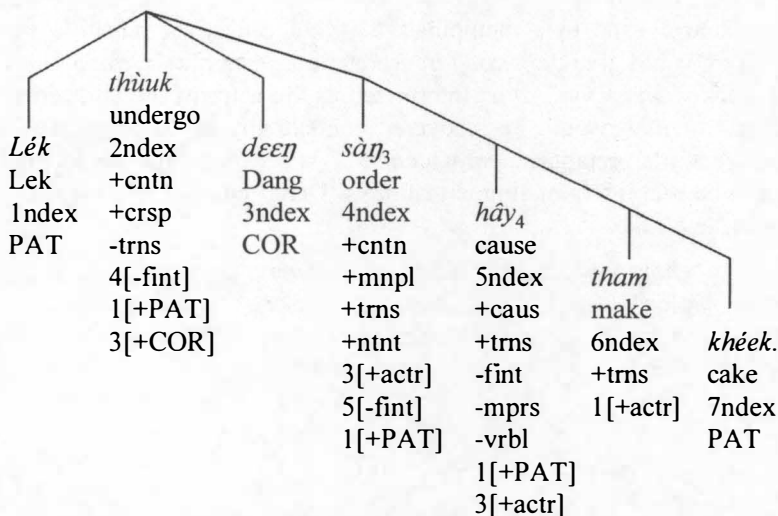
Dang ordered Lek to make a cake. (lit. Dang ordered Lek to cause (Lek) to make a cake.)

(64)b. *Lék nà? dɛɛŋ sàŋ₃ hây tham khéek.*

Lek TOP Dang order cause make cake

As for Lek, Dang ordered (her) to make a cake. (lit. Lek, Dang ordered to cause to make a cake.)

(64)c.



Lek was ordered by Dang to make a cake.

The Patient of the embedded *hây*₄ is always coreferential with the higher accusative Patient of the transitive manipulative verbal causative verb, in accordance with the P2P Control Rule demonstrated earlier in (44). Thus in (64a), the missing Patient of *hây*₄ must be coreferential with the higher Patient *lék* 'Lek'. *Dɛɛŋ* 'Dang', the closest available NP to the left, serves as its actor. The Regular Control Rule applies to the lower verb *tham* 'to make', since the higher verb *hây*₄ does not match the description of the P2P rule. Thus, *lék* 'Lek' is the actor of the lower verb. The fact that the noun phrase *lék* may be topicalised in example (64b) illustrates that the verb *sàŋ*₃ is transitive.

This P2P Control Rule still applies correctly in the related passive construction in (64c). First of all, by the chaining rule linking the upper Patient and lower Patient in the passive construction (see Appendix IV), *sàŋ*₃ 'to order' finds its missing Patient by copying the 1[+PAT] index from the regent verb *thiuk*. The actor of *sàŋ*₃ is *dεεŋ* 'Dang', the Correspondent actant of *thiuk*. Then, again in accordance with the P2P rule, the feature [?(+PAT)] in the lexical matrix of the embedded *hây*₄ copies the index of [1(+PAT)] from *sàŋ*. The other NP, *dεεŋ* 'Dang', is the actor of *hây*₄, since there are only two actants in this sentence and *dεεŋ* 'Dang' is again the closest available NP to the left, which is a dependent of the regent verb *thiuk*. As in (64a), *lεk* 'Lek', which is the actor of the lower verb *tham* 'to make', is identified as the higher Patient of its regent verb *hây*₄ by the Regular Actor Rule.

When the object of *hây*₄ is different from the patient of the higher verb, the sentence is unacceptable. As shown in (65), the object of *hây* is *dεεŋ* 'Dang', while the object of the higher verb *khǎw-sǎŋ*₂ 'to plead' is *phǎw* 'father'. The sentence hence is ungrammatical because it violates the P2P rule.

- (65) **Púk khǎw-sǎŋ₂ phǎw hây₄ dεεŋ pay thiaw.*
 Pook plead father cause Dang go travel
 AGT +trns PAT +trns PAT
 Pook pleaded with her father to allow Dang to go out.

Unlike the case of manipulative verbal causative intransitive verbs, the application of the Regular Actor Control Rule to a manipulative verbal causative transitive verb such as *khǎw-sǎŋ*₂ in (66a) would give the wrong interpretation. Such a case would allow *chán* 'I', which is the object of *khǎw-sǎŋ*₂, to be interpreted as the actor of the embedded *hây*₄. The unspecified object of *hây*₄ would be recovered contextually by the external linking rule, yielding an incorrect interpretation in which *chán* 'I' is causing some unspecified person to make a cake. The correct interpretation is that *dεεŋ* 'Dang' pleaded with *chán* 'I' and hence causes *chán* to make a cake.

- (66) a. *Dεεŋ khǎw-sǎŋ₂ chán hây₄ tham khéek.*
 Dang plead I cause make cake
 1ndex 2ndex 3ndex 4ndex 5ndex 6ndex
 AGT +cntn PAT +trns +trns
 +mnpl -fint -fint
 +trns *3[+actr] *x[+actr]
 3[+PAT] *x[+PAT]
 4[-fint]

*Dang pleaded with me to cause somebody to make a cake. (lit.)

- (66) b. *Dεεŋ khǎw-sǎŋ₂ chán hây₄ tham khéek.*
 Dang plead I cause make cake
 1ndex 2ndex 3ndex 4ndex 5ndex 6ndex
 AGT +cntn PAT +trns +trns
 +mnpl -fint -fint
 +trns 1[+actr] 3[+actr]
 3[+PAT] 3[+PAT] 6[+PAT]
 4[-fint]

Dang pleaded with me to make a cake. (lit. Dang pleaded with me to cause (me) to make a cake.)

To get the proper rule to apply for the correct interpretation, we need the convention that a more specific rule (the P2P rule) applies before a more general rule (the P2a rule). That is, if a form is eligible to undergo either the Regular Control Rule or the P2P rule, the latter, which specifically applies to continuing verbs, should apply first.

In (66b), the application of the P2P rule yields the correct coindexing. The missing Patient of *hây* is coreferential with the upper Patient *chán* 'I' according to the P2P Control Rule. The actor of *hây*₄, on the other hand, is *dɛɛŋ* 'Dang', the closest available NP to the left of *hây*. The actor of the lower verb *tham* 'to make' is *chán*, the Patient of *hây*₄, in accordance with the Regular Actor Control Rule.

4.1.2 PERSONAL CAUSATIVE VERB *hây* EMBEDDED UNDER MANIPULATIVE RESULTATIVE VERBS

*Hây*₄ 'to cause, to make' may appear to be embedded under manipulative resultative verbs which designate a change in the state of the Patient. Verbs of this class may be divided into two subclasses: correspondent manipulative resultative intransitive verbs, as in (67), and manipulative resultative transitive verbs, as in (68). The verb in the lower clause is a stative verb. However, it is questionable whether this *hây* is even a verb and, if so, whether it is the same as the causative verb *hây*₄.

- (67) *Kháw cà? kin₃ kúaytǎaw hây₄ kúaytǎaw môt.*
 he will eat noodles cause noodles use.up
 PAT +crsp COR
 +caus
 +mnpl
 -trns
 He will finish the noodles.

- (68) *Kháw cà? kin₄ kúaytǎaw hây₄ môt.*
 he will eat noodles cause use.up
 AGT +caus PAT
 +mnpl
 +trns
 He will finish the noodles.

Assuming that all verbs in Thai may appear as main verbs, if *hây* in (67) and (68) is a verb, then the embedded clause containing *hây* should be able to occur independently. Consider (69) and (70):

- (69) **Kháw hây kúaytiaw môt.*
 he cause noodles use.up
 He finished the noodles.
- (70) **Kháw hây môt.*
 he cause use.up
 He finished the noodles.

The forms *hây* in (69) and (70) may not appear as main verbs without context. Hence one might conclude that *hây* is not a verb in these constructions. However, this analysis in fact does consider *hây* in (67)–(68) to be a verb, not an adverb, for two reasons. First, *hây* in

(68) is similar to the causative *hây* discussed in the previous section in that it may take its own complement, namely *kuaytĭaw* 'noodles'. This is a major characteristic of a verb. An adverb, on the other hand, does not allow any dependent. Moreover, if *hây* were treated as an adverb, one would not be able to explain the presence of the second noun phrase *kuaytĭaw* in (67), which is also a dependent of the regent verb *kin*₃.

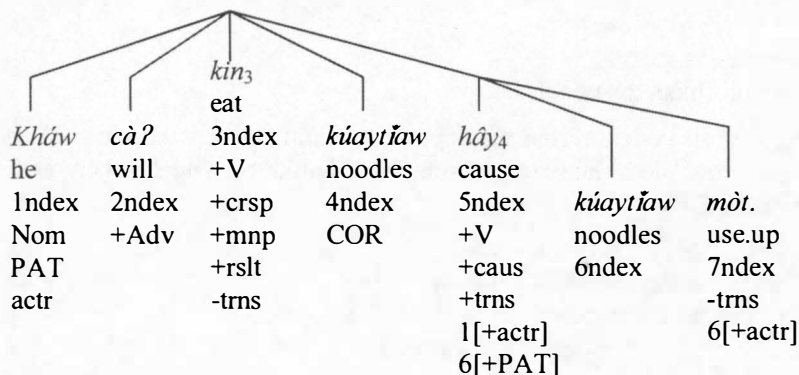
Second, the control rules for infinitival clauses, namely the upper and lower actor control rule (P2a) and the P2P rule, are applicable to the *hây* clause in these two constructions. These control chaining rules are applicable only to a verb and not to an adverb. As in the case of manipulative causative intransitive verbs, *hây* in (67) finds its missing actor from a higher Patient, following the P2a Control Rule. *Hây* in (68), on the other hand, allows its missing object to be coreferential with the Patient of a higher verb in accordance with the P2P rule, as in the case of manipulative causative transitive verbs. Moreover, since these rules are control rules for infinitival complements, this analysis suggests that the *hây* clauses in (67) and (68) are complements.

Because of the similarities between *hây* in (67) and (68) and the causative transitive *hây*₄ discussed in §4.1.1, this analysis concludes that *hây* in (67) and (68) and the causative verb *hây*₄ discussed in §4.1.1, are the same lexical entry, namely the personal causative transitive verb. This conclusion is determined by the ability of *hây* in (67) and (68) to take its own complement and to undergo the control rules for complements. However, the fact that *hây* in these constructions may not appear in a corresponding main verb position, while all other embedded verbs can, remains unexplained. Further study of this construction may shed some light on the matter.

In parallel to the analysis of manipulative verbal causative verbs, manipulative resultative verbs can be syntactically divided into two classes: correspondent manipulative resultative intransitive verbs and manipulative resultative transitive verbs, depending on which control rule is applied to each construction. The tree structures of (67) and (68) are shown in (67'a) and (68'a), respectively.

Correspondent Manipulative Resultative Intransitive Verbs

(67')a.



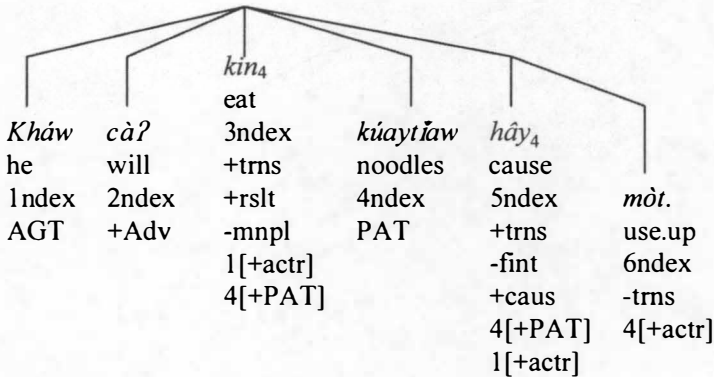
He will finish the noodles.

- b. **Kuaytĭaw nà? kháw cà? kin hây kuaytĭaw mòt.*
 noodles TOP he will eat cause noodles use.up
 As for noodles, he will finish them.

In (67'a), the verb *kin* is analysed as a corresponding manipulative resultative intransitive verb because its object *kúaytǎw* cannot be topicalised. The embedded verb *hây* finds its missing actor from the higher Patient *kháw* according to the Regular Actor Control Rule (P2a). Similarly, the lower verb *mòt* finds *kúaytǎw* as its higher Patient. Now consider the manipulative resultative transitive verb *kin* in (68'a) and (68'b).

Manipulative Resultative Transitive Verbs

(68'a).



He will finish the noodles.

- b. *Kúaytǎw nà? kháw cà? kin₄ hây₄ mòt.*
 noodles TOP he will eat cause use.up
 As for noodles, he will finish them.

Unlike *kin₃* in (67'a), *kin₄* in (68'a) can occur in a topicalised construction, as shown in (68'b). Therefore, this study considers *kin* in (68'a) to be a transitive verb. The embedded verb *hây* finds its missing Patient from the higher Patient *kúaytǎw* 'noodles', written as [4[+PAT]], in accordance with the P2P Control Rule. The NP *kháw* 'he' is interpreted as the actor of the causative verb *hây* because *kháw* 'he' is the closest available nominal to the left.

Now consider the *hây* clause in the following sentences.

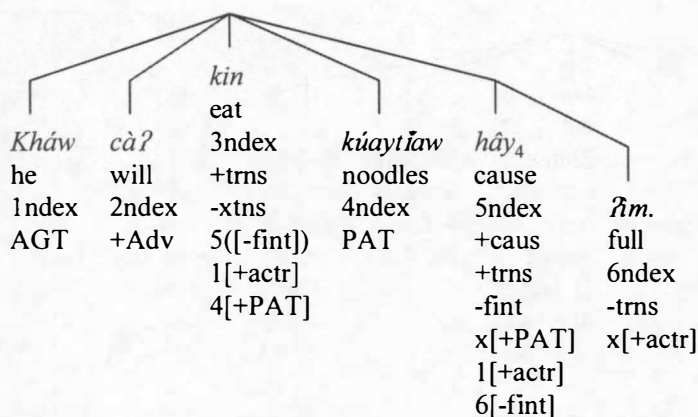
- (71) *Kháw cà? kin₁ kúaytǎw hây₄ ñm.*
 he will eat noodles cause full
He will eat the noodles until he is full.
- (72) *Kháw cà? kin₁ kúaytǎw hây₄ tua?εεɲ ñm.*
 he will eat noodles cause self full
He will eat the noodles until he is full.

The verb *kin₁* in (71) and (72) is a non-extension transitive verb. It does not require the presence of a dependent non-finite verb complement. Hence, the clauses introduced by *hây₄* in (71) and (72) are infinitival adjuncts. The actor of *hây₄* in (71) and (72) is interpreted as the higher actor *kháw* 'he' in accordance with the control rule for infinitival adjuncts.

The missing Patient of *hây₄* in (71) is a null anaphor recovered by the external linking rule in (56), which assigns the arbitrary index *x* to [?(+PAT)]. Grammatically, the *x* index which is written as 'xndex' could be interpreted as anyone; however, one cannot make anyone else full by eating food. Pragmatic considerations thus narrow the interpretation of the subject of *ñm* to *kháw*. Then if *kháw* is the actor of *ñm*, it must, by the Regular Actor

Control Rule, be coreferential to the object of *hây*. Since the actor and the Patient of *hây* are coreferential with the same word, the object can be expressed only as a reflexive. This interpretation is supported when the reflexive pronoun *tuaʔeeɣ* 'self' is present, as in (72). Example (71) is rewritten with the tree structure in (71'). The optionality of a non-finite clause introduced by *hây* is shown by the parentheses [5([-fint])] on the lexical matrix of *kin*₁.

(71')



He will eat the noodles until he is full.

An alternative analysis for *hây* in (67), (68), (71) and (72) could be that *hây* is an adverb, which would explain why *hây* may not appear as a main verb. In such an analysis, *hây* and the following noun in (67) and (72) could not form a constituent in Lexicase. However, the second *kúaytǎw* 'noodles' in (67) could then not occur as a sister of the head verb *kin* 'to eat', since that would make it a clausemate of a coreferential noun, the first *kúaytǎw*, violating the binding requirement that a noun must be free. The second *kúaytǎw* 'noodles' would instead have to be treated as the subject of the lower verb *mòt* 'use up'. That is, the lower verb would be a finite verb. However, the reflexive pronoun *tuaʔeeɣ* in (72) cannot be treated in the same way. A reflexive in Thai, as in most languages, needs an antecedent in the same clause; and therefore *tuaʔeeɣ* cannot be the subject of the lower verb *ʔm* 'to be full'. It must occur as a sister dependent of the verb *kin* 'to eat' in order to be bound with the subject *kháw* 'he'. The drawback of this alternative analysis is that not only does one need two different structures for the two sentences, but also one more often finds cognitive verbs such as *rúu* 'to know' and *bòk*₁ 'to tell', rather than action verbs such as *kin* 'to eat', followed by a finite verb. Thus, the verb analysis is preferable.

In short, I have demonstrated that there is a single personal causative transitive verb *hây*₄ which may appear in an embedded clause as a complement or as an adjunct. As in Kullavanijaya (1974), the analysis presented here treats the causative *hây*₄ and the ditransitive verbs *hây* 'to give' as distinct lexical entries because of their differences in meaning and distribution.

4.2 IMPERSONAL CAUSATIVE TRANSITIVE VERB *hây*₅

Impersonal verbs are verbs which do not allow referential subjects in their subcategorisation frame. In this paper *hây*₅ 'to let' is analysed as an extension impersonal causative transitive verb requiring a non-finite verb complement. This section discusses two

functions of the impersonal causative verb *hây*₅: as a marker for the third person imperative construction and as a complementiser.

4.2.1 IMPERSONAL CAUSATIVE TRANSITIVE VERB *hây*₅ FUNCTIONING AS AN IMPERATIVE FOR THE THIRD PERSON

Semantically, *hây*₅ 'to let' shows a speaker's indifference regarding the performance or non-performance of a third person's act. Consider (73a) and (73b).

- (73)a. *Hây*₅ *khruu maa ciŋciŋ thəʔ*.
 let teacher come true PRT
 +caus +V
 +mprs -fint
 +trns
 +xtns
 Let the teacher really show up.
- b. **Thəʔ hây*₅ *khruu maa ciŋciŋ thəʔ*.
 you let teacher come true PRT
 You let the teacher really show up.

*Hây*₅ in example (73a) implies indifference or defiance. Whether or not the action indicated in the *hây* clause takes place will have no effect on the speaker. The insertion of the second person pronoun subject *thəʔ* changes the grammaticality of the sentence, as illustrated in (73b). This incidence suggests that *hây*₅ in example (73a) is an impersonal causative verb. Now consider the same *hây*₅ in another example:

- (74) *Hây*₅ *kháw khəʔɔy pay siʔ*.
 let he wait thither PRT
 +trns
 +caus
 +mprs
 Let him wait!

Sentence (74) may have another interpretation in which *hây* is the personal causative transitive verb *hây*₄ (see §4.1). In this case, it allows a subject insertion, as shown in (75).

- (75) *Thəʔ hây*₄ *kháw khəʔɔy pay siʔ*.
 you cause he wait thither PRT
 +caus
 +trns
 -mprs
 You let him wait.

In other words, the impersonal causative *hây*₅ and the personal causative *hây*₄ differ only in that the latter allows a referential subject and the former does not. Another piece of supporting evidence for positing the impersonal causative *hây*₅ as a separate lexical entry is the fact that only *hây*₅, and no other Thai verb, allows third person imperatives. Consider (76).

- (76) *Khəɔy dǎaw náʔ.*
 wait soon how.about
 How about (you) waiting for a moment?

This sentence has only one interpretation: a speaker asks the hearer to wait. There is no interpretation such that the speaker would ask the second person to cause a third party to wait. The difference between *hây*, in (74) and (75), and other verbs, leads us to the conclusion that *hây*₅ in (74) is lexically marked as impersonal and deserves a separate lexical entry.

4.2.2 IMPERSONAL CAUSATIVE VERB *hây* FUNCTIONING AS A COMPLEMENTISER

*Hây*₅ ‘to let’ may introduce a clause in complex sentences, which suggests that this *hây* is a conditional complementiser.

- (77) a. *Hây₅ khruu maa ciŋciŋ chán kǝ́ɔ́ mây klua.*
 let teacher come really I also not fear
 Given the teacher really showing up, I am not afraid.
- b. **Thəə hây₅ khruu maa ciŋciŋ chán kǝ́ɔ́ mây klua.*
 you let teacher come true I also not fear
 Given the teacher really showing up, I am not afraid.

This analysis considers the form *hây* in (77a) to be the impersonal verb *hây*₅ functioning as a complementiser, not as a personal verb or a preposition. This decision is based on the following arguments. As with a verb in general, *hây* in (77a) may appear as a free verb in a root causative sentence, as shown earlier in (73a). Moreover, *hây* in (77a) does not allow the insertion of a subject, as shown in the ungrammatical (77b).

Note that (77b) would be acceptable if the form *hây* were the personal causative transitive verb *hây*₄ ‘cause’ cooccurring with the complementiser *méewáa* ‘even if’, as shown in (78).

- (78) *Méewáa thəə hây₄ khruu maa ciŋciŋ chán kǝ́ɔ́ mây klua.*
 even.if you cause teacher come true I also not fear.
 Even if you really make the teacher show up, I am not afraid.

The contrast in the ability of (77a) and (78) to allow the presence of a subject shows that the form *hây* in (77) and in (78) are two separate lexical items. While *hây*₄ in (78) is a personal verb, *hây*₅ in (77) is an impersonal verb functioning as a complementiser, introducing a concessive clause in a complex sentence.

Due to this parallel in interpretation between (73) and (77), the form *hây* in these two constructions is considered to be the same impersonal causative *hây*₅. This semantic evidence however, does not provide conclusive evidence that *hây* in (77a) is an impersonal verb. An alternative analysis might be that it is a preposition functioning as a complementiser. The preposition *méewáa* ‘even if’ in (79), like *hây* in (77a), allows the main clause to take a consequence aspect such as *kǝ́ɔ́* ‘also’. Since *méewáa*, which occurs in the same syntactic position, is a preposition, by analogy one might analyse *hây* as a preposition. The possibility that *hây* in constructions such as (77a) might better be analysed as a preposition is left open for further study.

- (79) *Méewâa khruu maa ciŋciŋ chán kɔ̌o mây klua.*
 even.if teacher come true I also not fear
 Even if the teacher shows up, I am not afraid.

5 ADVERB *hây*₆

This section will investigate the form *hây*₆ which appears in the frame:

NP V [+trns] NP ____ (NP)

in which the regent verb is a transitive verb and the noun following *hây* may be omitted contextually, as shown in the (a) and (b) examples of (80) and (81).

- (80) a. *Nídaa thǎi krapǎw hây₆ wiinaa.*
 Nida carry bag for Weena
 Nida carried a bag for Weena.

- b. *Nídaa thǎi krapǎw hây₆.*
 Nida carry bag for
 Nida carried a bag for (someone).

- (81) a. *Nídaa khǎay krapǎw hây₆ deɛŋ.*
 Nida sell bag to Dang
 Nida sold a bag to Dang.

- b. *Nídaa khǎay krapǎw hây₆.*
 Nida sell bag to
 Nida sold a bag to (someone).

Hây in (80) carries a benefactive meaning 'for'. In (81), *hây* is interpreted as a direction towards a goal, which may be equivalent to the English 'to'. In the following sections, different tests are applied to determine the status of the form *hây*₆.

5.1 TESTS FOR THE SYNTACTIC CATEGORY OF *hây*₆

Various tests may be applied to determine the status of the form *hây*₆, which may appear in the frame: NP V NP ____ NP, in comparison with the characteristics of a verb, a preposition, and an adverb. To begin, let us examine verbs, prepositions, and an adverb representing each category. The two verbs taken as verb models are *tòk* 'to fall' and *cèek* 'to distribute', as shown in (82) and (83).

- (82) *Kháw plàk dèk tòk lǔm nán.*
 he push child fall pit that
 He pushed a child down into that pit.

- (83) *Deɛŋ sǎi₂ khanǒm cèek dèkdèk.*
 Dang buy sweets distribute children
 Dang bought sweets to distribute to the children.

Two of the least controversial prepositions, *kàp* 'with' and *phǎa* 'for', have been selected to represent the characteristics of prepositions, as shown in (84) and (85). These two prepositions synchronically do not have corresponding verbs in the language.

- (84) *Deeg khay ríag níi kàp lék.*
 Dang talk story this with Lek
 Dang talked about this matter with Lek.
- (85) *Phố thamgaan nàk phía lúuk.*
 father work heavy for offspring
 A father works hard for his children.

Most adverbs in Thai are not followed by a noun. However, there exists an adverb *wáy* 'lying', which may or may not be followed by a locational noun, as shown in (86a) and (86b).

- (86) a. *Kháw thíg nágsǎi wáy.*
 he abandon book lying
 He left a book.
- b. *Kháw thíg nágsǎi wáy bâan nán.*
 he abandon book lying house that
 +lctn
 He left a book at that house.

The status of the word *wáy* 'lying' is controversial. It could be argued that *wáy* is a preposition. However, I consider *wáy* to be an adverb for the reason that it fails to show the characteristics of verbs and prepositions in the tests presented in this paper (to be illustrated in the next section). Therefore, I treat *wáy* as an adverb cooccurring with the regent verb *thíng* 'to abandon'.

The four tests used to examine the status of *hây* in (80a) and (81a) are: the free-verb test, the number-of-actions test, the topicalisation of *hây* together with the following NP, and the topicalisation of the NP after the form *hây*. Since *hây* in (80b) and (81b) is not followed by a noun, only the first two tests are applicable.

If *hây* in these constructions is a verb, it should be able to occur as a free verb while maintaining the meaning and syntactic restrictions it carried in (80) and (81). Furthermore, the NP after *hây* should be topicalisable, while *hây* together with the following NP should not. Moreover, there should be more than one action implied by the sentence. If *hây*₆ is a preposition, as claimed in previous analyses (Dejthamrong 1970, Kullavanijaya 1974), then the NP after *hây* should not be independently topicalisable, but rather *hây* with the following NP should form a PP constituent which can be topicalised. In addition, *hây* should not be stranded at the end of the sentence. If *hây*₆ is an adverb of the same class as

⁸ The adverb *cà?* test exemplified in (51) and (52) in §3 yields only a one-way implication: an element immediately dominated by *cà?* is a verb. However, not all verbs may cooccur with *cà?*. For example, the presence of *cà?* before the verb *tham* is unacceptable.

**supriya hây chán cà? tham khanǎm.*
 Supriya cause I will make sweets
 +V

Supriya made me make sweets.

Since *hây* in (80) and (81) does not allow the presence of *cà?*, this test is not a reliable criterion for determining the verbal status of *hây* in these constructions.

wáy 'lying', *hây* together with the following NP should not be topicalisable, since they do not form a PP constituent, while the NP after *hây* should be topicalisable (see §§5.1.1–5.1.4 for discussion). A summary of the results of these tests is shown in Table 1.

TABLE 1: THE RESULT OF THE TESTS FOR THE STATUS OF *hây*

	Free Verb	Number of Actions	Phrase Topicalisation	Toicalisation of NP
Preposition	-	1	+	-
Verb	+	2	-	+
Adverb <i>wáy</i>	-	1	-	+
<i>hây</i> in (80a)	-	1	-	+
<i>hây</i> in (80a)	-	1	-	+
<i>hây</i> in (81b)	-	1	N/A	N/A
<i>hây</i> in (81b)	-	1	N/A	N/A

The summary table shows that the free-verb test and the number-of-actions test agree that *hây* in (80) and (81) could be either an adverb or a preposition but not a verb. *Hây* with the following NP cannot be topicalised, as in the situation with a verb or an adverb, while the NP after *hây* may be topicalised, as with the NP occurring after a verb and after the adverb *wáy*.

5.1.1 THE FREE-VERB TEST

Only a verb, but not a preposition or an adverb, may appear as a free verb representing a valid sentence.

- (87) *Dèk tòk lǔm nán.*
child fall pit that
A child fell into that pit.
- (88) *Dɛɛŋ cɛ̀ɛk dèkdèk.*
Dang distribute children
Dang distributed (something) to the children.
- (89) **Dɛɛŋ káp lék.*
Dang with Lek
Dang with Lek.
- (90) **Phǎw phǎa lúuk.*
father for offspring
Father for children.
- (91) **Kháy wáy bǎan nán.*
he lying house that
He lying at that house.

Since *wáy* may not appear as a free verb in (91), this test supports the proposition that *wáy* is not a verb. If *hây* in the (a) and (b) examples of (80) and (81) is a verb, it should be able to appear as a free verb independently while preserving the meaning of *hây* in (80) and (81).

- (92) a. *Nidaa hây₆ wiinaa.*
 Nida for Weena
 *Nida for Weena.
- b. *Nidaa hây₆.*
 Nida for
 *Nida for.
- (93) a. *Nidaa hây₆ dɛɛŋ.*
 Nida to Dang
 *Nida to Dang.
- b. *Nidaa hây₆.*
 Nida to
 *Nida to.

The fact that *hây₆* in (92) and (93) cannot maintain the benefactive meaning ‘for’ or the directional meaning ‘to’ when appearing as a main verb indicates that *hây* in (80) and (81) is not a verb. *Hây* in these constructions is acceptable only when interpreted as ‘to give’.

5.1.2 THE NUMBER-OF-ACTIONS TEST

Li and Thompson (1973:176) have claimed that a sentence which contains more than one action should contain more than one verb. Although such subjective tests are not always reliable, it is apparent to Thai speakers that only (82) and (83) represent two actions, since both sentences contain a pair of verbs, *plàk* ‘to push’ and *tòk* ‘to fall’, and *s#₂* ‘to buy’ and *cêek* ‘to distribute’, respectively. On the other hand, sentences (84) and (85), containing the prepositions *kàp* ‘with’ and *pháa* ‘for’, and sentence (86), containing the adverb *wáy* ‘lying’, express only one action. If *hây* in (80) and (81) is a verb, the sentences should denote more than one action. In these examples, to carry a bag for someone in (80) and to sell a bag to someone in (81) reflect only one action. Thus, the number-of-actions test provides a piece of supporting evidence that *hây* in (80) and (81) is not a verb.

5.1.3 THE TOPICALISATION OF *hây₆* TOGETHER WITH THE FOLLOWING NP

While prepositional phrases may be topicalised, as in (94) and (95), verbs plus their objects may not be topicalised together, as shown in (96) and (97).

- (94) *Kàp lék nà? dɛɛŋ khəɲ khuy riŋ nɪi.*
 with Lek TOP Dang ever talk story this
 With Lek, Dang talked about this matter. (lit.)
- (95) *Pháa lúuk phɔ́w thamŋaan nàk.*
 for offspring father work heavy
 For their children fathers work hard. (lit.)
- (96) **Tòk lum nán kháw plàk dèk.*
 fall pit that he push child
 Into that pit, he pushed a child down. (lit.)

- (97) **Cèek dèkdèk dɛɛŋ sɨ̌₂ khanǽm.*
 distribute children Dang buy sweets
 To distribute to the children, Dang bought sweets. (lit.)

Consider the topicalisation of *wáy* together with the following NP in (98).

- (98) **Wáy bâan nán kháw thǐŋ náŋsǎi.*
 lying house that he abandon book
 At that house, he left a book. (lit.)

Unlike a prepositional phrase, *wáy* with the following NP cannot be topicalised together. This supports our claim that *wáy* is not a preposition. Similarly, the unsuccessful topicalisation of *hây* plus the following NP in (99) and (100) show that *hây* in (80) and (81) does not have the characteristic of a preposition in Thai. This test shows that *hây* behaves like a verb and like an adverb.

- (99) **Hây₆ wiinaa nídaa thǎi krapǎw.*
 for Weena Nida carry bag
 For Weena, Nida carried the bag.
- (100) **Hây₆ dɛɛŋ nídaa khǎay krapǎw.*
 to Dang Nida sell bag
 To Dang, Nida sold the bag.

5.1.4 THE TOPICALISATION OF NP AFTER *hây₆*

Examples (101) and (102) show that prepositions in Thai cannot be stranded and do not allow the following NP to be topicalised.

- (101) **Lék nà? dɛɛŋ khuy rǐaŋ níi kàp bǝy bǝy.*
 Lek TOP Dang talk story this with often
 As for Lek, Dang often talks with (her) about this matter.
- (102) **Lúuk nà? phǎw thamŋaan nàk phǎa.*
 offspring TOP father work heavy for
 As for children, fathers work hard for (them).

Examples (103), (104), and (105) show that NP occurring after verbs and the adverb *wáy* 'lying' may be topicalised. Both the verbs and the adverb *wáy* may be stranded.

- (103) **Lǔm nán nà? kháw khǎy plàk dèk tòk.*
 pit that TOP he ever push child fall
 As for that pit, he once pushed a child down (into it).
- (104) **Dèkdèk nà? dɛɛŋ sɨ̌₂ khanǽm cèek.*
 children TOP Dang buy sweets distribute
 As for the children, Dang bought sweets to distribute (to them).
- (105) **Bâan nán nà? kháw thǐŋ náŋsǎi wáy.*
 house that TOP he abandon book lying
 As for that house, he left a book (there).

Unlike NP after prepositions, the NP after *hây₆* may be topicalised, similar to the NP after verbs and the adverb *wáy*.

- (106) *Wiinaa nà? nídaa thǎi krapǎw hây₆.*
 Weena TOP Nida carry bag for
 As for Weena, Nida carried the bag for (her).
- (107) *Dɛɛŋ nà? nídaa khǎay krapǎw hây₆.*
 Dang TOP Nida sell bag to
 As for Dang, Nida sold the bag to (her).

5.2 ANALYSIS OF *hây₆*

The results of the four tests agree that *hây* in (80) and (81) behaves like the adverb *wáy*, rather than like a verb or a preposition. Verbs and prepositions are counterindicated by two tests. The free-verb test and the semantic test show that *hây* in (80) and (81) is not a verb. Furthermore, the topicalisation of the NP after *hây₆* and the impossibility of topicalising *hây₆* with the following NP rule out the possibility that *hây₆* is a preposition.

The present analysis considers *hây* in both the (a) and (b) examples of (80) and (81) to belong to a single class of adverb, for the following reasons. First, since verbs in Thai have the potential to occur as free verbs, that is as main verbs of independent clauses, the inability of *hây₆* to appear as a free verb, demonstrated in (92) and (93), provides a strong piece of evidence that *hây* in (80) and (81) is not a verb. This claim is supported by the number-of-actions test. Second, in Lexicase, a preposition always needs a dependent cohead in an exocentric construction. If *hây₆* were a preposition, then *hây₆* together with the following NP should form a PP constituent and permit topicalisation. In contrast, examples (99) and (100) illustrate that *hây* and the following noun phrase cannot be topicalised together. Third, although a preposition cannot be stranded, as shown in (101) and (102), *hây* in (80) and (81) may be left stranded, as shown in (106) and (107), indicating that *hây* is not a preposition. Finally, all of the four tests demonstrate that *hây* exhibits the same characteristics as the adverb *wáy*.

Furthermore, our claim that *hây* in (80) and (81) is an adverb also allows us to explain why the NP after *hây* may be left out when the context is given, in contrast with the NP after a preposition in Thai, which may not be omitted. This claim is further supported by the analysis of the benefactive *gei* 'for' in Chinese, *cho* 'for' in Vietnamese, and *hir* 'for' in Tai Nung, which have the corresponding verb 'to give' and which have been analysed as derived adverbs when they appear without the presence of a following noun (Starosta (1985:224) for Chinese, and Clark (1992:146–147) for Vietnamese and Tai Nung).

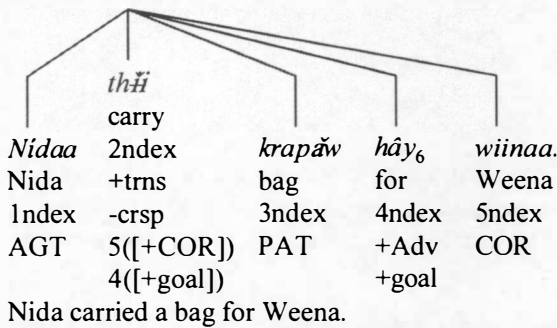
Hây in (80) and in (81) shows similarities in distribution, as discussed. I shall claim further that the form *hây* in these two constructions belongs to a single lexical entry *hây₆*, despite the differences in interpretation. The difference in interpretation between (80) and (81) is governed by the difference in the classes of the regent verbs. When *hây₆* appears with correspondent ditransitive verbs such as *khǎay₂* 'to sell', *sǎw₂* 'to teach', and *mǎw₂* 'to

⁹ This is analogous to the analysis of case inflection systems. The same case inflection may encode different meanings, depending on the verb or preposition with which it cooccurs, and no one would propose setting up distinct case inflection categories for different functions.

deliver', it is interpreted as a direction towards a goal, corresponding to the English 'to'. Conversely, when it appears with non-correspondent transitive verbs such as *sĥĭ* 'to buy', *thĥĭ* 'to carry', *khăy₁* 'to sell', and *sĥm₁* 'to teach', it carries the meaning of a benefactive action towards a goal and is interpreted as 'for'. Thus the adverb *hây₆* is represented by the localistic feature [+goal]. Moreover, as illustrated in (99) and (100), *hây* and the following noun phrase do not form a constituent which can be topicalised together. This analysis hence considers the two elements to be grammatically independent of each other.

Since *hây₆* behaves similarly in the two examples for each group (80) and (81), I will discuss only the constructions (80a) and (81a) in this section. Illustration (108) shows the tree structure of the clause containing the non-correspondent transitive verb *thĥĭ* 'carry'. *Thĥĭ* allows but does not require the presence of the adverb *hây* 'for' and the following correspondent noun *wiinaa* 'Weena', shown by the parentheses [4([+goal])] and [5([+COR])] on the governing verb *thĥĭ*. That is, *hây* and *wiinaa* are both considered to be independent adjuncts of *thĥĭ*.

(108)



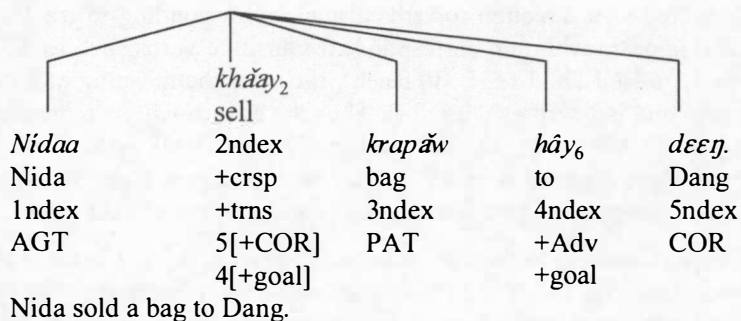
Unlike other adjuncts, *hây₆* 'for' is needed to disambiguate a benefactive interpretation from a possessive interpretation. *Hây₆* cannot be omitted if the benefactive interpretation is to be maintained, as illustrated in (109). In other words, *hây₆* in (80a) and (108) is obligatory due to semantic interference.

- (109) *Nidaa thĥĭ krapăw wiinaa.*
 Nida carry bag Weena
 *Nida carried a bag for Weena.
 Nida carried Weena's bag.

The tree structure of a sentence containing the correspondent ditransitive verb *khăy₂* 'to sell', which expects the cooccurrence of *hây₆*, is shown in (110). In other words, *hây* and the following NP in this structure are complements of *khăy₂* 'to sell'.¹⁰

¹⁰ Besides this correspondent ditransitive verb *khăy₂*, there is another correspondent ditransitive *khăy₃*, which may cooccur with the preposition *kĕe* 'to' and an optional adverb *hây₆* indicating goal, as shown in the following examples (a) and (b).

(110)



There are two pieces of evidence supporting the claim that *hây*₆ and the following NP cooccurring with correspondent ditransitive verbs are complements, while *hây* and the following NP cooccurring with non-correspondent transitive verbs are adjuncts.

First, the goal adverb *hây*₆ cannot appear with the non-correspondent verb *s#*₁ 'to buy' to give the interpretation 'to'. This is shown in the unacceptable interpretation of (111).

- (111) *Nidaa yip*₁ *krapăw hây*₆ *deεŋ*.
 Nida pick.up bag to Dang
 *Nida picked up a bag to Dang.

Second, only the directional *hây*₆ may occur closer to the head, as in (112). Sentence (113) shows that the first *hây* can only be interpreted as the goal direction and not as the benefactive. This analysis suggests that the directional *hây*₆ is a complement, while the benefactive *hây*₆ is an adjunct.

- (112) *Chán mōɔp dōɔkmáay chō nán hây*₆ *khruuyà y hây*₆ *khruu léεw*.
 I deliver flower bunch that to principal for teacher already
 I have already delivered that bouquet of flowers to the principal for the teacher.
- (113) *Chán mōɔp dōɔkmáay chō nán hây*₆ *khruu hây*₆ *khruuyà y léεw*.
 I deliver flower bunch that teacher for principal already
 *I have already delivered that bouquet of flowers for the teacher to the principal.
 I have already delivered that bouquet of flowers to the teacher for the principal.

Furthermore, the goal adverb *hây*₆ may occur with non-correspondent non-causative affect transitive verbs, such as *dū?* 'to reproach', and *tii* 'to hit', as in (114). In this case, the

-
- a. *Nidaa khăăy*₃ *krapăw hây*₆ *kēε chán*.
 Nida sell bag to to I
 1ndex 2ndex 3ndex 4ndex 5ndex 6ndex
 AGT +trns +Adv +P COR
 +crsp +goal +goal
 5[+goal]
 6[+COR]
 4([+goal])
 Nida sold a bag to me.
- b. *Nidaa khăăy*₃ *krapăw kēε chán*.
 Nida sell bag to I
 Nida sold a bag to me.

sentence has a malefactive interpretation. The form *hây* does not precede a noun, and can be omitted, as shown in (115).

- (114) *Đãw chán cà? tii thəə hây₆*
 soon I will hit you
 Any minute now, I will hit you.

- (115) *Đãw chán cà? tii thəə*
 soon I will hit you
 Any minute now, I will hit you.

When *hây* in (114) is followed by a noun, it is interpreted as benefactive, as in (116).

- (116) *Đãw chán cà? tii thəə hây₆ mēē.*
 soon I will hit you for mother
 For mother's sake, I will hit you any minute now.

Because of the complementary distribution between the interpretation of the form *hây* in (114) and (116), the form *hây* in both sentences is considered to be the same lexical item *hây₆*. The differences in meaning may be attributed to pragmatic usage of the verb in each sentence. The goal adverb *hây₆* is interpreted as malefactive when it occurs with verbs which are perceived as unfavourable. Such verbs are, for example, *kàt* 'to bite', *tii* 'to hit', *dũ?* 'to reproach' *dàa* 'to scold', *kròot* 'to be angry', and *yēēŋ* 'to snatch'. *Hây₆* may also occur with verbs which do not carry unfavourable meanings by themselves, such as *cùup* 'to kiss', if the action is perceived as threatening or destructive.

In short, the form *hây* in (80) and (81) is shown to be a single word, namely an adverb *hây₆* indicating a goal. This *hây₆* may appear with non-correspondent transitive verbs to carry the benefactive meaning, as in (80) and (116), or the malefactive meaning, as in (114), in which case it marks an adjunct. When *hây₆* cooccurs with a correspondent ditransitive verb, it carries a directional meaning and is a complement. *Hây₆* cooccurring with correspondent ditransitive verbs may carry the benefactive meaning only if there is an additional form *hây* bearing the directional interpretation.

Claiming that *hây₆* cooccurring with an optional noun is an adverb is somewhat counterintuitive when seen from the perspective of English grammar. However, as shown in earlier tests, the form *hây* in (80) and (81) fails to exhibit the prominent characteristics of prepositions and verbs. Thus, the adverb analysis is the most preferable one, linguistically. It is, however, possible that this adverb *hây₆* is derived from prepositions (Clark: pers. comm.). An alternative analysis which assumes that *hây* could be a transitive preposition when it is followed by a noun and an intransitive preposition when the following noun is not present (see Emonds 1976:172) is ruled out here, because to formalise such an analysis within the constrained theory of Lexicase is not possible.

6. CONCLUSION

This study concludes that there are altogether six different homophonous forms of *hây* in Thai. There are three homophonous ditransitive verbs *hây*: *hây₁* requires two bare noun phrases as complements; *hây₂* takes a bare noun phrase and a prepositional phrase as complements; *hây₃* requires two bare noun phrases and a verb complement. The two causative verbs, namely the personal causative verb *hây₄* and the impersonal causative verb

hây₅, differ in their ability to allow a subject. Moreover, *hây₅* implies the speaker's indifference to or defiance of the action of a third person. Finally, *hây₆*, of the benefactive meaning 'for' and of the directional meaning 'to', is considered to be a single lexical entry of the adverb category because it fails to exhibit the characteristics of either verbs or prepositions, a finding which contrasts with Dejthamrong (1970) and Thepkarnchana (1986). Differences in interpretation in the different uses of *hây₆* are governed by the different classes of the regent verbs.

APPENDIX I: ABBREVIATIONS

Acc	Accusative	lctn	location	PAT	Patient
actr	actor	lit	literal translation	PP	prepositional phrase
Adv	adverb	LOC	Locus	PRT	particle
afct	affect	mnpl	manipulative	rslt	resultative
AGT	Agent	mprs	impersonal	TOP	topic marker
caus	causative	N	noun	trns	transitive
cntn	continuing	ndex	index of	V	verb
COR	Correspondent	Nom	Nominative	vrbl	verbal
crsp	correspondent	NP	noun phrase	xtns	extension
fint	finite	P	preposition		

APPENDIX II: CONTINUING VERBS IN THAI

Continuing verbs are verbs which are interpreted by the Patient-to-Patient Control Rule (P2P). The following lists exemplify members of continuing verbs in Thai, based on Indrambarya's (1994) analysis.

1. Correspondent Continuing Intransitive verbs:

thiuk₅ 'to undergo' *doon₄* 'to undergo'

2. Non-manner Continuing Non-manipulative Transitive Verbs:

hàn₃ 'to chop' *sàŋ₆* 'to order (food)' *yìp₃* 'to pick up'
tii₅ 'to hit' *sĕĕ₃* 'to buy' *yé₂* 'to sew'

3. Non-manner Continuing Manipulative Transitive Verbs:

a. Verbal causative continuing transitive verbs

chǎn₅ 'to invite' *khǎwǎŋ₃* 'to plead' *waan₄* 'to ask'
sàŋ₇ 'to order' *bǎwǎ₇* 'to order' *cháy₇* 'to order'
plàk₅ 'to push' *nĕŋnam₄* 'to suggest' *chuan₅* 'to persuade'

b. Resultative Continuing Transitive Verbs

tii₆ 'to hit' *khían₆* 'to whip' *yìk* 'to pinch'
sĕĕ₄ 'to buy' *kin₄* 'to eat' *thǎŋ* 'to memorise'

4. Correspondent Non-benefactive Transitive Verbs:

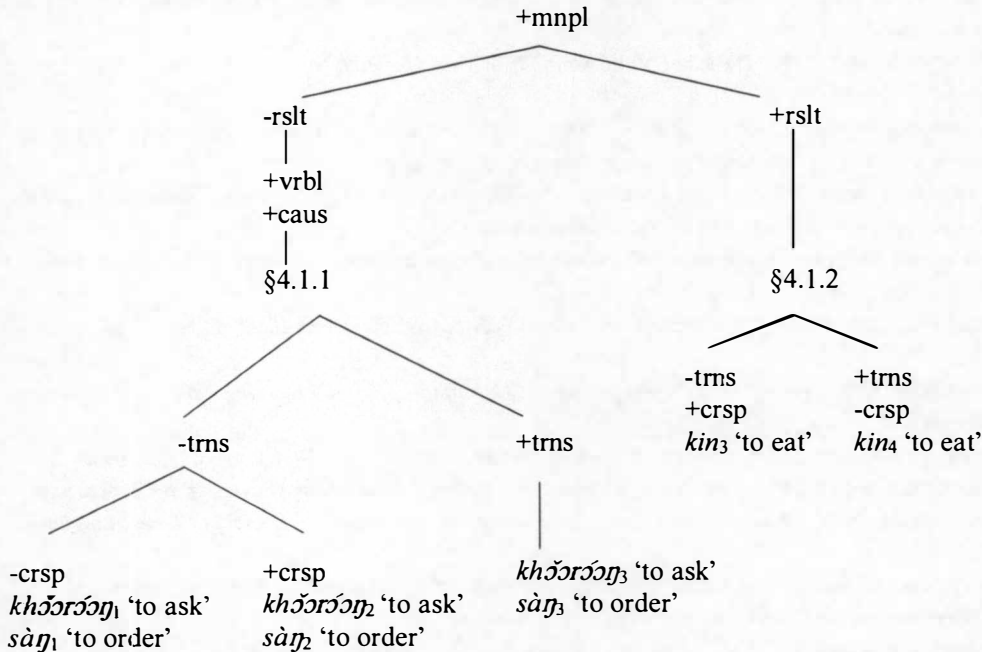
hây₃ 'to give'

5. Correspondent Benefactive Transitive Verbs:

*cêk*₅ 'to distribute' *hât* 'to show' *lâw* 'to relate'
pôn 'to feed to somebody's mouth'

APPENDIX III: SUBCATEGORISATION OF MANIPULATIVE VERBS

Manipulative verbs are verbs which require the presence of the causative verb *hây* in the embedded complement clause. The following table illustrates the subcategorisation of manipulative verbs which can be semantically divided into verbal causative manipulative and resultative manipulative verbs



APPENDIX IV: CONTROL CHAINING RULES

1. Regular Actor Control Rule:

$$\left| \begin{array}{c} ?[+actr] \\ -fint \\ nndex \end{array} \right| \quad \rightarrow \quad [m[+actr]] \quad \backslash \quad \left| \begin{array}{c} m[+PAT] \\ n[-fint] \end{array} \right|$$

(Starosta 1992:47)

2. PAT-to-PAT Control Rule (P2P):

$$a. \quad \left| \begin{array}{c} -fint \\ ?[+PAT] \\ nndex \end{array} \right| \quad \rightarrow \quad [m[+PAT]] \quad \backslash \quad \left| \begin{array}{c} +cntn \\ m[+PAT] \\ n[-fint] \end{array} \right|$$

b. The actor of the non-finite verb is interpreted as the closest available noun-headed dependent of a regent verb to the left of the embedded verb (Indrambarya 1994:300).

REFERENCES

- Chomsky, Noam, 1981, *Lectures on Government and Binding: the Pisa lectures*. Dordrecht: Foris.
- Cole, Peter, 1987, Null objects in Universal Grammar. *Linguistic Inquiry* 18/4:597–612.
- Clark, Marybeth, 1992, Serialisation in mainland Southeast Asia. *Pan-Asiatic linguistics: Proceedings of the Third International Symposium on Language and Linguistics*, vol. I:145–159. Bangkok: Chulalongkorn University.
- Dejthamrong, Orathai, 1970, Grammatical functions of the word *hây* in the Thai language. MA thesis, Chulalongkorn University, Bangkok.
- Emonds, Joseph, 1976, *A transformational approach to English syntax: root, structure-preserving, and local transformations*. New York: Academic Press.
- Hasagawa, Yoko, 1988, Question pull: a diagnostic test for the complement and adjunct distinction in Japanese. *Berkeley Linguistic Society* 13:66–77.
- Indrambarya, Kitima, 1990, The status of the word *hây* in Thai. *University of Hawai'i Working Papers in Linguistics* 22/1:33–71.
- 1994, Subcategorisation of verbs in Thai: a lexicase dependency approach. Doctoral dissertation, University of Hawai'i, Honolulu. Ann Arbor: University Microfilms International.
- Kullavanijaya, Pranee, 1974, Transitive verbs in Thai. Doctoral dissertation, University of Hawai'i, Honolulu. Ann Arbor: University Microfilms International.
- Li, Charles and Sandra A. Thompson, 1973, Coverbs in Mandarin Chinese: verbs or prepositions? *Journal of Chinese Linguistics* 2/3:257–278.
- Noss, Richard, 1964, *Thai reference grammar*. Foreign Service Institute, Department of State, Washington D.C.
- Pagotto, Louise, 1985, On impersonal verbs in English. *University of Hawai'i Working Papers in Linguistics* 17/2:1–69.
- Panupong, Vichin, 1970, *Inter-sentence relations in modern conversational Thai*. Bangkok: the Siam Society. (Reprinted from Chantavibulya, Vichin, 1962. Doctoral dissertation, University of London.)
- Starosta, Stanley, 1985, Mandarin case marking: a localistic lexicase analysis. *Journal of Chinese Linguistics* 13/2:216–265.
- 1988, *The case for lexicase: an outline of lexicase grammatical theory*. London: Pinter Publishers.
- 1992, Lexicase revisited: a revision and extension of 'lexicase'. MS.
- Thepkarnchana, Kingkarn, 1986, Serial verb constructions in Thai. Doctoral dissertation, University of Michigan. Ann Arbor: University Microfilms International.

VARIABILITY IN CAMBODIAN COPULAR CONSTRUCTIONS: A SEMANTIC ANALYSIS

CHARLES PAUS

1. INTRODUCTION¹

1.1 CAMBODIAN 'TO BE'

Cambodian, like most of the world's languages, does not have a single element which corresponds to all of the senses of English 'to be': rather, as many as seven different forms are used in existential, locative, and copulative constructions. In general, each Cambodian construction corresponds to a distinct part of the range of functions and meanings of 'to be', although some constructions appear to overlap partially with others. The basic inventory of Cambodian forms of 'to be' can be summarised as follows:

1) existential: *mian*

2) locative: (*kii*) *nəw*

3) copulative

a) with predicate adjectives: zero, *kii*

b) in equative constructions with predicate NPs:

kii – sometimes exclusively, sometimes in variation with the less preferred options *ciə*, *kii ciə* (formal style) and zero (informal style)

c) in non-equative constructions with predicate NPs: *ciə* and *kii* usually judged equally acceptable, but *ciə* sometimes preferred; *kii ciə* (formal style) and zero (informal style) always possible but less preferred options

1.2 PREVIOUS ANALYSES

I am not aware of any detailed treatment of this issue in the literature on Cambodian. Some discussion can be found in basic grammars such as Ehrman (1972), Jacob (1968), and Noss and Proum (1966). However, the descriptions offered in these grammars are incomplete, and they differ on some important points, especially where copulative constructions are concerned.

For example, in Ehrman, no semantic distinction is made between the copulas *ciə*, *kii*, *kii ciə* and the zero copula. The copula *ciə* is defined as a copula which indicates that "the subject is a member of the class of objects named by the predicate"; it is the copula used in sentences such as 'This is a pencil' (1972:75). Concerning *kii*, Ehrman says that it "may be used with

¹ I have been assisted in my fieldwork on Cambodian by Mr Van So Chau, a native speaker of Cambodian.

the same meaning as *ciə*, except that where it is followed by more than one proper name, it is obligatory"; an example which requires *kii* is 'This is Sok, Suon, and Sos' (p.75). No examples are given for *kii ciə*, but Ehrman claims that it is "used like *ciə*" (p.76). Finally, the use of the zero copula is said to occur "in some cases...especially in predicates containing time words"; for example, in 'Today is Saturday' (p.76).

Jacob (1968:140–141) suggests that *ciə* and *kii* are semantically distinct. She claims that "*ciə* 'to be'...may be used when two nouns are linked by it", and she gives examples such as 'I am a teacher' and 'That girl is Mr. X's daughter' (p.140). Concerning *kii*, she notes that it is used "like *ciə* in linking two nouns but has the lexical meaning 'to be in essence, to be by nature'", as in the example 'What's this? It's coral' (p.141). Finally, she points out that two noun phrases may be linked by zero where *ciə* or *kii* would be expected; these are regarded as "ad hoc verbalisations of nominal constructs" as in the example 'She's a widow' (p.141).

Noss and Proum (1966:118–119) offer the most detailed account of which I am aware. They draw clear semantic distinctions between *ciə*, *kii*, and the zero copula. The word *ciə* is defined as 'to be a member of the class of', as in 'I'm a student' (p.119). In contrast, *kii* is defined as 'to be equivalent to', as in 'The biggest book of all is the one on the bottom' (p.118). Finally, the zero copula is interpreted as 'to be characterised by', as in 'My book is white' (p.118). In addition, Noss and Proum discuss the use of *nəw* as the locative 'to be (somewhere)', as in 'The blue book is on the bottom' (p.118), and the use of *mian* as the existential 'to be (in existence)', as in 'There are two books' (p.118).

Of the three descriptions, the one offered by Noss and Proum seems to be the best; in addition to being the most detailed, it is the one which, in general, most closely matches the material which I have elicited in fieldwork sessions with Mr Van So Chau. However, the match between Noss and Proum's description and the system exhibited by Mr Chau is not always exact. In addition, the overlap of functions of some variants, which is mentioned in Ehrman and Jacob but not in Noss and Proum, is abundant in my elicited material.

2. THE FORMS OF 'TO BE'

In the following sections, the various Cambodian forms and uses of 'to be' as given by Mr Chau are discussed individually.

2.1 EXISTENTIAL 'TO BE': *mian*

The verb *mian* is used for existence (EXT): 'there is/are', 'to be in existence'. No other copula is possible.

- (1) *Mian siəwphəw bəy.*
EXT book three
There are three books.

The same verb is used for possession ('to have'); the possessor NP precedes the verb:

- (2) *Kñom mian siəwphəw bəy.*
I EXT book three
I have three books.

2.2 LOCATIVE 'TO BE': (*kii*) *nəw*

The Cambodian verb *nəw* means 'to be located' and is used to link a subject to a predicate indicating temporary or permanent location in space or time. In all contexts, *nəw* can be optionally preceded by the equative *kii*; there is no change in meaning, but the expanded form is characterised by Mr Chau as "unusual" and likely to occur in formal rather than informal styles. Locatives (LOC) are not formed with *mian*, *ciə*, *kii ciə*, or zero.

Examples (3)–(6) illustrate the use of (*kii*) *nəw* to indicate, respectively, temporary location of non-events in space, permanent location of non-events in space, spatial location of events and temporal location of events:

- (3) *Tidaa (kii) nəw pnum.piñ.*
 Tidaa COP LOC Phnom.Penh
 Tidaa is in Phnom Penh.
- (4) *Pnum.piñ (kii) nəw kmər.*
 Phnom.Penh COP LOC Cambodia
 Phnom Penh is in Cambodia.
- (5) *Tii pracuŋ (kii) nəw / kii / ?ciə sabaa.*
 place meeting COP LOC COP COP school
 The meeting is at school.
- (6) *Peil pracuŋ (kii) nəw / kii / ?ciə maon prampii.*
 time meeting COP LOC COP COP hour seven
 The meeting is at seven o'clock.

In examples (5) and (6), the copulas *kii* and, more marginally, *ciə* can be used in place of *nəw* to link a subject NP with a predicate NP. In Cambodian, these constructions must be regarded as copulative rather than locative, although this distinction is lost in the English translations. The use of copulas in such constructions is treated in detail below, in the discussion of Cambodian copulative constructions.

2.3 COPULATIVE CONSTRUCTIONS WITH STATIVE PREDICATES: zero, *kii*

Statives in Cambodian, unlike adjectives in English, are full verbs and so do not require any other accompanying verb, such as a copula (COP). It may be noted that *nəw*, *mian*, *ciə*, and *kii ciə* are never possible in stative constructions. The behaviour of *kii* is unique; while it is nearly always less preferred than zero, it is "less bad" than the other variants in combination with a stative verb:

- (7) *Tidaa (?kii) lʔaa.*
 Tidaa COP good
 Tidaa is good.
- (8) *Tidaa (?kii) lʔaa krup peil.ve.liə.*
 Tidaa COP good all time
 Tidaa is good all the time.

(10) Viə (?kii) lʔaa dael tidaa səmlap trəy.
it COP good that Tidaa kill fish
It's good that Tidaa killed the fish.

(11) Dael tidaa səmlap trəy kii lʔaa
 ?
 that Tidaa kill fish COP good
 That Tidaa killed the fish is good.

2.4 IDENTIFICATION AND ATTRIBUTION: COPULATIVE CONSTRUCTIONS WITH NP PREDICATES: *kii*, *ciə*, *kii ciə*, zero

In general, *kii* is an equative copula, whose function is to identify one NP with another; the basic meaning of *A kii B* is something like '(the thing named by) A is equal to (the thing named by) B'. In contrast, *ciə* is a copula with attributive function; it links a subject NP with a predication. The basic meaning of *A ciə B* is something like '(the thing named by) A can be characterised as B' or 'B is a characteristic of (the thing named by) A'.

(12) *Giscard d'Estaing is the president of France.*

In the equative reading, both NPs function as referential expressions; each refers to an individual, and the copula simply indicates that there is identity between the two referents. In the predicative reading, the predicate NP does not have a referential function; it does not refer directly to some individual, but rather tells something about the subject of the sentence. In the predicative reading, 'is the president of France' describes the referent of the subject NP in the same way that predicates such as 'likes to play tennis', 'was born in France' or 'is six feet tall' give information about the referent of the subject NP.

The distinction between predicative and equative copulas is one that is of little importance in English, since the same copula is used for both functions. However, the distinction can be shown even in English; when (12) has a predicative reading, the article can be optionally omitted. When it has an equative reading, the article cannot be omitted.

In Italian, the distinction between these two types of copula sentences has clearer syntactic consequences. Compare the following (from Moro 1991:21):

- (13) *Una foto del muro fu la causa della rivolta.*
A picture of the wall was the cause of the revolution.

The predicate NP of this sentence can be replaced by the clitic *lo*, which precedes the copula:

- (14) *Una foto del muro lo fu.*
A picture of the wall was it.

However, if the order of the NPs in the original sentence is reversed, then clitic replacement of the predicate NP is not possible:

- (15) *La causa della rivolta fu una foto del muro.*
The cause of the revolution was a picture of the wall.
(16) **La causa della rivolta lo fu.*
The cause of the revolution was it.

To explain the difference between (13) and (15), Moro (p.22) appeals to the notion of referentiality versus non-referentiality of NPs. A referential NP "points to" a referent, while a non-referential NP describes another NP. In Italian, a predicate NP can be replaced by the clitic *lo* only if the predicate NP is non-referential. In (13), clitic replacement is possible because 'the cause' is non-referential; 'a picture' can be assigned the property of being 'the cause'. A property is a predication and can be replaced by a clitic. In (15), clitic replacement is not possible because 'a picture' is referential; 'the cause' cannot be assigned the property of being 'a picture'.

The distinction which Moro makes between copulative constructions with a non-referential predicate NP and copulative constructions with a referential predicate NP corresponds to the distinction made by Lyons between predicative and equative copulative constructions. Perhaps instead of the terms non-referentiality/referentiality and predicative/equative, the terms 'attribution' and 'non-attributive identification' might be used to distinguish these two types of copulative constructions.

The distribution of *kɨi* and *ciə* in Cambodian can be analysed in terms of the distinction between these two types of copulative constructions. In equative constructions, in which the predicate NP is referential, or identificational rather than attributive, *kɨi* is preferred, sometimes exclusively, sometimes with *ciə* as a less preferred option. In predicative

The word *kii* is used exclusively in sentences with a clearly equative sense, such as (17)–(20) below, in which the copula can be translated as ‘is equal to’. The order of constituents can be changed without changing the meaning of these sentences, or the choice of copula; the copula functions here much like an equals sign:

- Equative sentences are not limited to mathematical statements of equivalence; in (21) and (22), the copula does not introduce a predication, but simply indicates identity of subject and predicate. Only *kii* can be used in these sentences:

- [illegible]

- (24) *Kñom kii viə.*
 *ciə
 I COP it
 I am it.

When the subject NP is formally identical to the predicate NP, as in (25), the sense of the construction is one of identification rather than predication, and *kii* is the only possible copula:

- (25) *Kñom kii kñom.*
 *ciə
 I COP I
 I am I.

Exclusive use of *kii* is often found where the head of the subject NP is formally identical to the head of the predicate NP, although the subject NP and the predicate NP may have different modifiers:

- (26) *Siəwphəw thom.ciən.kɛ kii siəwphəw pǝa khiew.*
 *ciə
 book biggest COP book colour blue
 The biggest book is the blue book.
- (27) *Siəwphəw pǝa khiew kii siəwphəw thom.ciən.kɛ*
 *ciə
 book colour blue COP book biggest
 The blue book is the biggest book.

The same holds for constructions with *?aa* 'one' as the head of the predicate NP:

- (28) *Siəwphəw thom.ciən.kɛ kii ?aa pǝa khiew.*
 *ciə
 book biggest COP one colour blue
 The biggest book is the blue one.
- (29) *Siəwphəw pǝa khiew kii ?aa thom.ciən.kɛ*
 *ciə
 book colour blue COP one biggest
 The blue book is the biggest one.

To summarise, constructions in which *kii* is the only possible copula seem to be limited to constructions which have a purely equative or identificational sense, and which retain the sense of equation or identification (and exclusive use of *kii*) when the order of constituents is reversed.

2.4.2 PREFERENCE FOR *kii*

In the majority of copulative sentences with a predicate NP in my elicited material, what we have is either relative preference for *kii*, with *ciə* as a less preferred option, or else *kii* and *ciə* are judged equally acceptable. The relative degree of preference for *kii* can be changed by changing the order of constituents. For the variants which correspond to the Italian sentence in (13) (in which clitic replacement is possible in Italian), with an attributive (or 'non-referential') predicate NP, *kii* and *ciə* are judged equally acceptable. For the variants

Here are several minimal pairs of this type:

- The copulative variants of examples (5) and (6), given above in the section on *nəw* and repeated here as (38) and (39), follow the same pattern. When the predicate NP is one of identification rather than attribution, as in these examples, then *kii* is preferred over *ciə*.

However, when the order of constituents is reversed, the predicate NP is attributive, and *kii* and *ciə* are equally acceptable:

- (38) *Tii pracuŋ (kii) nəw / kii / ?ciə sabaa.*
 place meeting COP LOC COP COP school
 The meeting is at school.
- (39) *Peil pracuŋ (kii) nəw / kii / ?ciə maon prampii.*
 time meeting COP LOC COP COP hour seven
 The meeting is at seven o'clock.
- (40) *Sabaa kii tii pracuŋ.*
ciə
 school COP place meeting
 School is the place of the meeting.
- (41) *Maon prampii kii peil pracuŋ.*
ciə
 hour seven COP time meeting
 Seven o'clock is the time of the meeting.

Examples (42) and (43) suggest that pronouns are more 'referential' than proper names in Cambodian:

- (42) *Niən kii tidaa.*
ciə
 she COP Tidaa
 She is Tidaa.
- (43) *Tidaa kii niən.*
?ciə
 Tidaa COP she
 Tidaa is she.

Examples (44) and (45), which follow the same pattern, suggest that a first name is more 'referential' than a title in Cambodian:

- (44) *Tidaa kii nea?.srəy kuit.*
ciə
 Tidaa COP Mrs Kuit
 Tidaa is Mrs Kuit.
- (45) *Nea?.srəy kuit kii tidaa.*
?ciə
 Mrs Kuit COP Tidaa
 Mrs Kuit is Tidaa.

2.4.3 DEMONSTRATIVES

Demonstratives appear to be inherently 'referential' in Cambodian; this seems logical, since demonstratives have the inherent function of 'pointing to' a referent. In non-interrogative constructions, when either the subject NP or the predicate NP contains a demonstrative, *kii* is preferred over *ciə*:

- (46) *Nih kii mdaay robah kmeyn.srəy.*
 ?ciə
 this COP mother of girl
 This is the mother of the girl.
- (47) *Mdaay robah kmeyn.srəy kii nih.*
 ?ciə
 mother of girl COP this
 The mother of the girl is this one.
- (48) *Srəy nih kii tidaa.*
 ?ciə
 woman this COP Tidaa
 This woman is Tidaa.
- (49) *Tidaa kii srəy nih.*
 ?ciə
 Tidaa COP woman this
 Tidaa is this woman.

2.4.4 WH-QUESTIONS

Similarly, interrogative constructions with the WH-element *nah* 'which' inherently ask for identification rather than description; *kii* is always preferred over *ciə* in such constructions:

- (50) *Muy nah kii trəy?*
 ?ciə
 one which COP fish
 Which is a fish?
- (51) *trəy muy nah kii trəy l?aa?*
 ?ciə
 fish one which COP fish good
 Which fish is a good fish?

The behaviour of the WH-elements *wəy* 'what' and *nea?naa* 'who' is exceptional. In constructions with *wəy*, *kii* and *ciə* are judged equally acceptable, whether the sense of the construction is one of asking for identification, as in (52), or one of asking for a definition or description, as in (54) (example (53) can have either reading):

- (52) *Nih kii wəy?*
 ciə
 this COP what
 What is this?
- (53) *Viə kii wəy?*
 ciə
 it COP what
 What is it?

- (54) *Baay kii wəy?*
 ciə
 rice COP what
 What is rice?

In constructions with *nea?naa* 'which person, who', *ciə* is always preferred over *kii*:

- (55) *Nih ciə nea?naa?*
 ?kii
 this COP who
 Who is this?
- (56) *Niən ciə nea?naa?*
 ?kii
 she COP who
 Who is she?
- (57) *Tidaa ciə nea?naa?*
 ?kii
 Tidaa COP who
 Who is Tidaa?

2.4.5 PREFERENCE FOR *ciə*

Other than the 'who' sentences, I was able to elicit only a few sentences for which Mr Chau expressed a preference for *ciə*, with *kii* as a secondary option. My data contain no sentences in which *ciə* is chosen exclusively; wherever *ciə* is possible, *kii* is available, at least as a secondary option.

Sometimes, there is a correspondence between preference for *ciə* and the use of the indefinite article with the predicate NP in the English translation. The English translations for (58) and (59) differ only in that (58) has an indefinite article where (59) has a definite article; (58) has preference for *ciə*, and (59) has preference for *kii*:

- (58) *Viə ciə siəwphəw l?aa.*
 ?kii
 it COP book good
 It is a good book.
- (59) *Viə kii siəwphəw l?aa.*
 ?ciə
 it COP book good
 It is the good book.

However, this correspondence does not always obtain; compare example (32) above, repeated here as (60), in which *kii* and *ciə* are judged equally acceptable, whether the English translation has a definite article or an indefinite article as part of the predicate NP:

- (60) *Tidaa kii nea?.kruu knuŋ phum nih.*
 ciə
 Tidaa COP teacher in village this.
 Tidaa is the/a teacher in the village.

The opposition of identification and attribution seems to play a more important role in the distinction between preference for *ciə* and preference for *kii* than the distinction between definiteness and indefiniteness of predicate NPs. Sentence (58), with preference for *ciə*, is a likely response to a question which asks for a description, such as 'What kind of book is it?'. Sentence (59), with preference for *kii*, is a likely response to a question which asks for identification, such as 'Which book is it?'.

Similarly, *ciə* is preferred in (61) below, a response to the question 'Can you tell me something about Tidaa?', while *kii* is preferred in (62), a response to the question 'Who is Tidaa?':

- (61) *Tidaa ciə nea?.kruu.*
 ?*kii*
 Tidaa COP teacher
 Tidaa is a teacher.

- (62) *Tidaa kii nea?.kruu.*
 ?*ciə*
 Tidaa COP teacher
 Tidaa is a teacher.

I was surprised by the judgments given for two minimal pairs which seem structurally equivalent to examples (30) to (37). For those sentences, the variants with an identificational NP have preference for *kii*, and the variants with an attributive predicate NP have *kii* and *ciə* judged equally acceptable. In the following examples, the variants with an identificational predicate NP also have preference for *kii*, but the variants with an attributive predicate NP have preference for *ciə*:

- (63) *Rud.tiəni kmər kii pnum.piñ.*
 ?*ciə*
 capital Cambodia COP Phnom.Penh
 The capital of Cambodia is Phnom Penh.

- (64) *Pnum.piñ ciə rud.tiəni kmər.*
 ?*kii*
 Phnom.Penh COP capital Cambodia
 Phnom Penh is the capital of Cambodia.

- (65) *Pratiniətapəday kii George.Bush.*
 ?*ciə*
 president COP George.Bush
 The president is George Bush.

- (66) *George.Bush ciə pratiniətapəday.*
 ?*kii*
 George.Bush COP president
 George Bush is the president.

Mr Chau explained that he prefers *ciə* in (64) and (66) because "everyone knows that these are so, so there's no need to identify them". In other words, the sense of (64) and (66) is felt to be strongly descriptive or attributive, and not likely to be a question of identification. It seems then that, at least in some cases, there may be an element of subjective judgment involved in the use of *kii* and *ciə*, which is related to how the speaker perceives the relative

plausibility of identificational and attributive readings. The fact that *kii* and *ciə* are judged equally acceptable in structurally similar examples such as (30) to (37) may be due to the fact that for these sentences, identificational and attributive readings are interpreted as equally reasonable.

In conclusion, it appears that despite the high degree of overlap in the distribution of *kii* and *ciə*, the two copulas have essentially distinct functions; *kii* functions primarily as an equative copula, linking a subject NP with a referential, or non-attributive, identificational predicate NP, while *ciə* functions primarily as a predicative copula, linking a subject NP with a non-referential, or attributive, predicate NP. At least some of the overlap of the two copulas in my data appears to result from the fact that many sentences permit both equative and predicative readings, especially when elicited in isolation, in the absence of any particular context. In a future study of Cambodian copulative constructions, the hypotheses which I have presented here might be confirmed or refined through examination of the use of the copulas as they occur in particular contexts in Cambodian discourse or written texts.

REFERENCES

- Ehrman, Madeline, 1972, *Contemporary Cambodian: a grammatical sketch*. Washington: Foreign Service Institute.
- Jacob, Judith, 1968, *Introduction to Cambodian*. London: Oxford University Press.
- Lyons, John, 1977, *Semantics*, vol.1. London: Cambridge University Press.
- Moro, Andrea, 1991, Referentiality and non-referentiality in Italian copula sentences. MS, University of Venice.
- Noss, Robert and Im Proum, 1966, *Cambodian basic course*, vol.1. Washington: U.S. Department of State.

